Digital Object Identifier XXX

Pattern or Not? QAOA Parameter Heuristics and Potentials of Parsimony

VINCENT EICHENSEHER[1], MAJA FRANZ[1], CHRISTIAN WOLFF[2], WOLFGANG MAUERER[1,3]
[1]Technical University of Applied Sciences Regensburg, Regensburg, Germany
[2]University of Regensburg, Regensburg, Germany
[3]Siemens AG, Foundational Technologies, Munich, Germany

Corresponding authors: Vincent Eichenseher (vincent.eichenseher@othr.de), Maja Franz (maja.franz@othr.de)

ABSTRACT Structured variational quantum algorithms such as the Quantum Approximate Optimisation Algorithm (QAOA) have emerged as leading candidates for exploiting advantages of near-term quantum hardware. They interlace classical computation, in particular optimisation of variational parameters, with quantum-specific routines, and combine problem-specific advantages—sometimes even provable—with adaptability to the constraints of noisy, intermediate-scale quantum (NISQ) devices. While circuit depth can be parametrically increased and is known to improve performance in an ideal (noiseless) setting, on realistic hardware greater depth exacerbates noise: The overall quality of results depends critically on both, variational parameters and circuit depth. Although identifying optimal parameters is NP-hard, prior work has suggested that they may exhibit regular, predictable patterns for increasingly deep circuits and depending on the studied class of problems.

In this work, we systematically investigate the role of classical parameters in QAOA performance through extensive numerical simulations and suggest a simple, yet effective heuristic scheme to find good parameters for low-depth circuits. Our results demonstrate that: (i) optimal parameters often deviate substantially from expected patterns; (ii) QAOA performance becomes progressively less sensitive to specific parameter choices as depth increases; and (iii) iterative component-wise fixing performs on par with, and at shallow depth may even outperform, several established parameter-selection strategies. We identify conditions under which structured parameter patterns emerge, and when deviations from the patterns warrant further consideration. These insights for low-depth circuits may inform more robust pathways to harnessing QAOA in realistic quantum compute scenarios.

INDEX TERMS Combinatorial Optimisation, Heuristics, Parametrisation, QAOA, Quantum Computing

I. INTRODUCTION

Despite progress towards error-corrected quantum computers [1], noise and imperfections remain dominant in NISQ devices [3, 16]. Moreover, algorithms with proven quantum advantage [23] are impractical at current qubit scales. Over the past decade, research has centred on hybrid quantum–classical algorithms adapted to restricted hardware. These methods replace full coherent quantum evolution with iterative classical protocols that invoke quantum subroutines, thereby reducing qubit and gate requirements [35, 37]. This mitigation limits noise accumulation on small-scale NISQ implementations. Such variational quantum algorithms (VQAs) [9] include the variational quantum eigensolver (VQE) [35] and the quantum approximate optimisation algorithm (QAOA) [11].

QAOA admits variable circuit depths by repeating parametrised mixer and cost unitaries, that is, the number of layers p. As the algorithm corresponds to a Trotterisation of adiabatic quantum evolution [3], the corresponding dynamics are recovered in the limit $p \to \infty$ and yields the optimal solution in the absence of noise.

Identifying optimal parameters for finite p is known to be NP-hard [4]. Nonetheless, for small and modest depths, both empirical and theoretical studies indicate that good approximations are feasible for many problems [3, 8, 39, 38, 17]. While classical optimisation remains the most common strategy [5, 13, 29], alternative approaches have been proposed that yield high-quality parameters at reasonable computational cost [42, 14, 10, 44, 41, 25].

Several studies suggest the existence of discernible pat-

terns in optimal QAOA parameters [49, 28, 32]. Typically, cost-unitary parameters are observed to increase smoothly with depth, while mixer-unitary parameters decrease. Empirical results further indicate that parameters at depth p differ only slightly from those at $p+1$ [49]. This behaviour aligns with the fact that larger p correspond to a more fine-grained Trotterisation of adiabatic time evolution. In this work, we re-examine the assumption that optimal parameters adhere to such patterns by systematically investigating the cost landscapes of optimisation-free parameter initialisation methods for the canonical NP-complete problems *MaxCut*, *VertexCover*, and *Max3SAT*. In particular, we examine linear ramp schedules [32] and compare their performance with parameters obtained through classical optimisation and a brute-force sequential method. Our analysis includes not only parameters consistent with the established patterns, but more importantly seeks systematic deviations, thereby accounting for degenerate optima and relating varying degrees of pattern-conformity to the expected solution quality. This enables an assessment of the extent to which optimal parameters realistically follow proposed patterns. In contrast, prior approaches typically focus solely on parameters identified by optimisation routines or other selection methods, offering no insight into the performance of neighbouring points in the optimisation landscape dynamics [2]. In summary, our main contributions are:

1) We replicate and confirm—using a comprehensive open source reproduction package—results reported in the literature, and further analyse the influence of component-wise parameter adjustments on corresponding cost-landscapes. We employ large-scale numerical simulations for deep circuits beyond 20 layers.
2) We examine the conformance of optimal QAOA parameters to patterns suggested in the literature, and show that actually optimal parameters often *fail* to follow these. We provide guidance on when such patterns are likely to hold, as well as when and to what degree deviations may arise. We identify possible reasons, and uncover more accurate patterns.
3) We propose a novel *sequential* method for QAOA parametrisation, and demonstrate that it yields, despite its structural simplicity and little required computational effort, comparatively good results akin to more complex schemes at low depths.

The rest of this article is structured as follows. Sec. II outlines the theoretical background and terminology. Related work is reviewed in Sec. III, followed by the experimental setup in Sec. IV. Results are presented in Sec. V, with attention to symmetries, parameter concentration, and fixed-parameter implications. Sec. VI considers the potential of optimisation-free approaches and the role of parameter patterns. We conclude in Sec. VII.

II. FOUNDATIONS

As QAOA [11] has been extensively studied and is well known, suffice it to summarise its essential features here. The heuristic approximates solutions to unconstrained optimisation problems by measuring the quantum state obtained after applying p layers of alternating unitaries to the initial state $|s\rangle = |+\rangle^{\otimes n}$.

Each layer $i \in [1, p]$ applies a phase-separation operator

$$H_C = \sum_{i,j} J_{ij}\sigma_z^{(i)}\sigma_z^{(j)} + \sum_i h_i\sigma_z^{(i)}, \tag{1}$$

where J_{ij} specifies interaction strengths and h_i denotes local energy offsets. Classical variables are encoded via $\sigma_z^{(i)}$. This is followed by a mixer unitary $U_B(\beta_i) = e^{-\imath\beta_i H_B}$ that commutes with the phase-separation operator. When $|s\rangle$ is a product of σ_x eigenstates, the transverse-field Hamiltonian $H_B = \sum_{i=0}^{n}\sigma_x^{(i)}$ is typically chosen for this purpose.

The repeated application of these layers results in the parametrised quantum state

$$|\gamma, \beta\rangle = U_B(\beta_p)U_C(\gamma_p)\cdots U_B(\beta_1)U_C(\gamma_1)\,|s\rangle\,, \tag{2}$$

that depends on $2p$ real parameters $\vec{\gamma} = (\gamma_1, \gamma_2, \cdots, \gamma_p)$ and $\vec{\beta} = (\beta_1, \beta_2, \cdots, \beta_p)$. Parameter optimisation seeks parameters $\vec{\gamma}, \vec{\beta}$ that minimise (or maximise) the expectation value

$$F_p(\vec{\gamma}, \vec{\beta}) = \langle\vec{\gamma}, \vec{\beta}|\, H_C\,|\vec{\gamma}, \vec{\beta}\rangle \tag{3}$$

that can be sampled from the quantum circuit implementing Eq. 3. Optimisation is usually performed by classical iterative methods that produce a sequence of gradually improving choices for $(\vec{\gamma}, \vec{\beta})$.

In general, $\vec{\gamma}$ is 2π-periodic and $\vec{\beta}$ is π-periodic. The expectation value is also invariant under time reversal, $F_p(\vec{\gamma}, \vec{\beta}) = F_p(-\vec{\gamma}, -\vec{\beta})$, and problem-specific symmetries may further restrict the parameter space. For MaxCut on regular graphs, $\beta \in [-\pi/4, \pi/4]$ and $\gamma \in [-\pi/2, \pi/2]$, whereas for non-regular graphs—or in general, other optimisation problems—only the general periodicity applies, so $\gamma \in [-\pi, \pi]$.

The expectation value in Eq. 3 can be decomposed into a sum of terms that involve only qubits i and j:

$$F_p(\vec{\gamma}, \vec{\beta}) = \sum_{i,j} J_{ij}\,\langle\vec{\gamma}, \vec{\beta}|\,\sigma_z^{(i)}\sigma_z^{(j)}\,|\vec{\gamma}, \vec{\beta}\rangle\,. \tag{4}$$

The collection of terms involving qubits i or j constitutes the *reverse causal cone* of their correlation function. For each depth p, this cone contains only finitely many terms [40], and its size depends solely on p, independent of the instance size n.

III. RELATED WORK

QAOA was introduced by Farhi *et al.* in 2014, and has since inspired numerous variants and extensions [18, 7, 47, 45, 42]. As simulating its output is classically intractable even for depth $p = 1$ [12] given some widely accepted complexity-theoretic assumptions, QAOA is frequently seen

as indicator of quantum advantage. While the single-layer case is well understood [25, 11, 7, 15], the behaviour of deeper circuits remains less clear, notwithstanding a steadily expanding body of results [49, 46, 48, 28, 27].

In particular, this includes the identification of patterns across problem instances [40, 33], for instance through machine learning [20], analytical considerations [25], or by other approaches [49, 28]. Such patterns aim to provide less computationally demanding alternatives than orthodox optimisation to determine good parameters $\vec{\beta}, \vec{\gamma}$. In general, standard approaches are costly and require multiple quantum circuit evaluations, although the exact effort depends on many details [34, 13, 43].

Zhou *et al.* [49] observed that optimal parameters exhibit smooth trends, with γ_i increasing and β_i decreasing as functions of the layer index $i \in [1, p]$. Exploiting these patterns, they proposed to let parameters obtained at low depth serve as initial values for deeper circuits, either via interpolation or through low-depth amplitudes in a frequency-domain approximation. Both approaches were shown to outperform random initialisation of QAOA. Brandão *et al.* [6] demonstrated that parameters concentrate for related problem instances (*e.g.*, MaxCut on three-regular graphs), with the variance of optimal parameters decreasing as instance size increases. This suggests that either sufficiently large instances or averages over many smaller instances are necessary to identify a meaningful concentration point. By formally considering infinite system size through restriction to the reverse causal cone, Streif and Leib [40] further showed that parameter concentration depends not on the global problem size n, but rather on qubit correlations.

Refs. [24, 33, 32] suggest that parameters initialised via linear ramp (LR) schedules, with γ increasing linearly and β decreasing linearly, generalise effectively across different problems and instances. Such schedules also provide robust initialisations for subsequent optimisation [33].

IV. METHOD

In the following we present our sequential parameter initialisation method. Furthermore, we outline the considered metrics and optimisation problems, as well as our experimental setup.

A. SEQUENTIAL PARAMETER INITIALISATION

As a baseline for assessing alternative optimisation and initialisation strategies, we propose a *sequential* parameter optimisation scheme: QAOA parameters (β, γ) are sampled from a uniform 32×32 grid over $[-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\pi, \pi]$, or, where symmetry permits, over $[-\frac{\pi}{4}, \frac{\pi}{4}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ (see Sec. V-A). The procedure begins at depth $p = 1$, with all 32^2 grid points evaluated. At each depth p, the parameters yielding the lowest energy expectation are fixed. The depth is then incremented to $p + 1$, and the same parameter grid is explored with previously fixed components held constant. This process, which is illustrated in Fig. 1, continues until the target depth p_{target} is reached. This way, only a restricted region of the optimisation landscape close to the chosen fixed parameters, requires evaluation, leading to linear rather than exponential scaling in the number of QAOA evaluations with depth p. Although conceptually simple, fixing the best available tuple (β, γ) at each stage may be suboptimal, as the effect of additional layers on the energy landscape remains uncertain.

The strategy of fixing lower-level QAOA parameters while iteratively increasing circuit depth has precedent, notably in Ref. [28]. There, parameter selection is guided by an optimiser rather than grid evaluation, and all components of (β, γ) are re-optimised at each iteration, in contrast to our approach, which fixes only the highest-depth components.

B. METRICS

We assess the performance of a given parameter set $(\vec{\gamma}, \vec{\beta})$ using the energy expectation value $F_p(\vec{\gamma}, \vec{\beta})$ (cf. Eq. 3). To enable comparison across problem instances, we employ the residual energy

$$r = \frac{F(\beta, \gamma) - E_0}{E_{\text{max}} - E_0}, \tag{5}$$

as metric, where E_0 and E_{max} denote the ground- and maximum-excited-state energies of the problem Hamiltonian H_C. For multiple instances, we compute the mean residual energy $\bar{r}$ and its standard deviation σ_r. While broadly consistent with other quality measures, such as the energy approximation ratio, the residual energy is particularly suited to evaluating states that are good but not optimal [40].

C. COMBINATORIAL PROBLEMS

We consider three seminal NP-complete problems: MaxCut, VertexCover and Max3SAT.

- **MaxCut** seeks a partition of a graph that maximises the number of crossing edges [22]. Our study considers three-regular graphs, consistent with the QAOA literature, where d-regular graphs are standard [49, 6, 13, 48, 46, 11]. We evaluated 40 instances in total: ten each of orders 10, 12, 14, and 16.
- **VertexCover** seeks a minimal set of nodes such that every edge is adjacent to at least one node in the set [22]. For our experiments, we employ the same three-regular graph instances as in the MaxCut study: ten each of orders 10, 12, 14, and 16.
- **Max3SAT** involves assigning Boolean variables in a conjunctive normal form (CNF) formula (with three Boolean variables per clause, 3-CNF) so as to maximise the number of satisfied clauses [22, 31]. Instances are characterised by the clause-to-variable ratio $\alpha = |\text{Clauses}|/|\text{Variables}|$. For small α, formulas are typically under-constrained with many solutions; for large α, they become over-constrained and unsatisfiable. Hard instances concentrate in $\alpha \in (3.5, 4.9]$, where satisfiability occurs with a probability of about 50%. The critical point shifts with formula size, converging

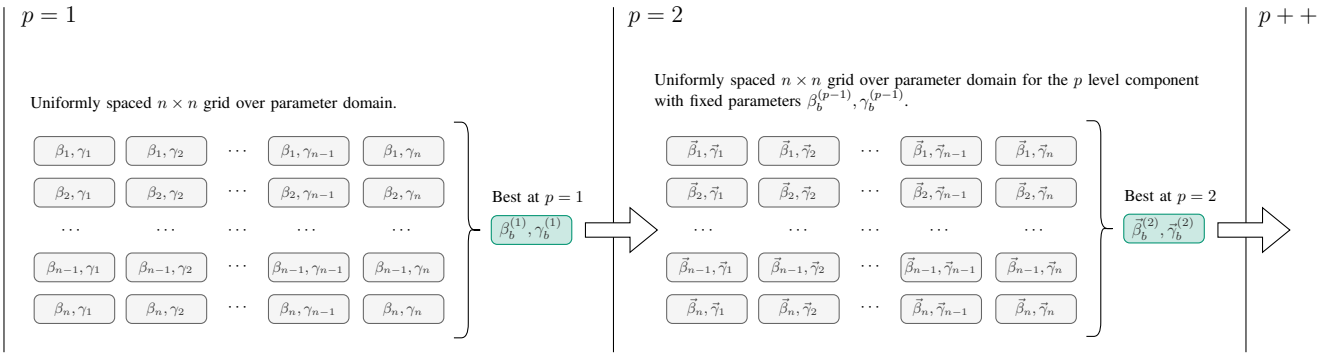

FIGURE 1: Proposed sequential parameter initialisation method.

near $\alpha \approx 4.25$ for large systems [31, 26]. In our study, instances with $\alpha \in (3.5, 4.9]$ are designated "hard", and those outside this range "easy". We evaluate ten instances each, comprising 14–24 qubits (a CNF formula with n variables and m clauses is represented by $n + m$ qubits).

D. EXPERIMENTAL SETUP

Additionally to the sequential method, we also explore the optimisation landscapes of the following parameter selection strategies:

- In the **LR-QAOA** scheme, parameters are set according to linear ramp (LR) schedules, following Montañez-Barrera and Michielsen [32]. The ramps are parameterised by circuit depth p and slopes $(\Delta\beta, \Delta\gamma)$, corresponding to the mixer and phase operators, respectively. Given differing implementations of mixers in the QAOA literature (*e.g.* Refs. [18, 19]), we consider two cases: (1) Both ramp parameters $\Delta\beta, \Delta\gamma$ are positive, denoted as $\text{LR}_{+\beta}$, and the evolution uses $U_B(\beta_j)$ as defined in Sec. II. (2) One parameter is negative $-\Delta\beta, \Delta\gamma$, denoted as $\text{LR}_{-\beta}$. This case is also applied in Ref. [32], likely as it aligns the mixer with the initial state $|x\rangle$. Previous work indicates that QAOA performs better when the initial state coincides with the mixer's ground state [19]. For both cases, we use $\Delta\gamma = 0.6$ and $|\Delta\beta| = 0.3$ as suggested by the authors.
- As a **classical optimiser**, we employ the gradient-free optimiser COBYLA [36] to optimise parameters. Initial values are obtained using sequential, $\text{LR}_{-\beta}$ and $\text{LR}_{+\beta}$ parameters for MaxCut, VertexCover and Max3SAT, respectively. These varied initialisations ensure sufficiently distinct starting points, allowing us to assess how the choice of initialisation influences optimiser convergence and final parameter quality.

To probe the parameter landscape of LR and classical optimisation methods, we first obtain parameters for the target depth p_{target} using the method of interest. Similar to the sequential method, for each $p \in [1, p_{\text{target}}]$, parameters from for the depth $p-1$ are fixed, while the parameters at level p are specified by a uniformly spaced grid. However, unlike in the sequential method, these grid evaluations merely serve as an assessment tool and do not influence the method's chosen parameters. This procedure enables an organised assessment of parameter quality and allows comparison of parameters given by the subject initialisation methods with nearby alternatives in the optimisation landscape. Further details on this landscape scan procedure are described in Appendix A.

For all problem instances, we set $p_{\text{target}} = 7$. Additionally, for MaxCut, we examine the $p_{\text{target}} = 21$ landscapes for the same instances, incrementing p by 2 in each iteration. When optimising the $p_{\text{target}} = 21$ instances, the parameters are initially set to the best values identified in the $\text{LR}_{+\beta}$ $p = 21$ landscape.

To eliminate detrimental influence of imperfections in actual quantum systems, all experiments were conducted using ideal simulations with Qiskit 1.0.2 [21]. The results are fully reproducible [30] via our reproduction package (link in PDF).

V. RESULTS

We now commence to discussing details of our numerical simulations, in particular with respect to properties of the parameter optimisation landscapes, as well as variations in and quality of results. Note that we postpone discussing interpretation and consequences to Sec. VI, and concentrate on the empirical observations in the following.

A. QAOA PARAMETER SYMMETRIES

Our initial set of experiments aims to provide an intuitive visual understanding of symmetries that are inherent to the properties of QAOA (cf. Sec. II), restrict the input domain and simplify numerical simulation. Restricted symmetries for MaxCut on 3-regular graphs and general symmetries are illustrated in Fig. 2.

The periodicity of the energy landscape can be observed in both parts of the figure, with period $\frac{\pi}{2}$ for β and π for γ, resulting in two horizontal identical sections in the general case; for the case of MaxCut this results in four horizontal

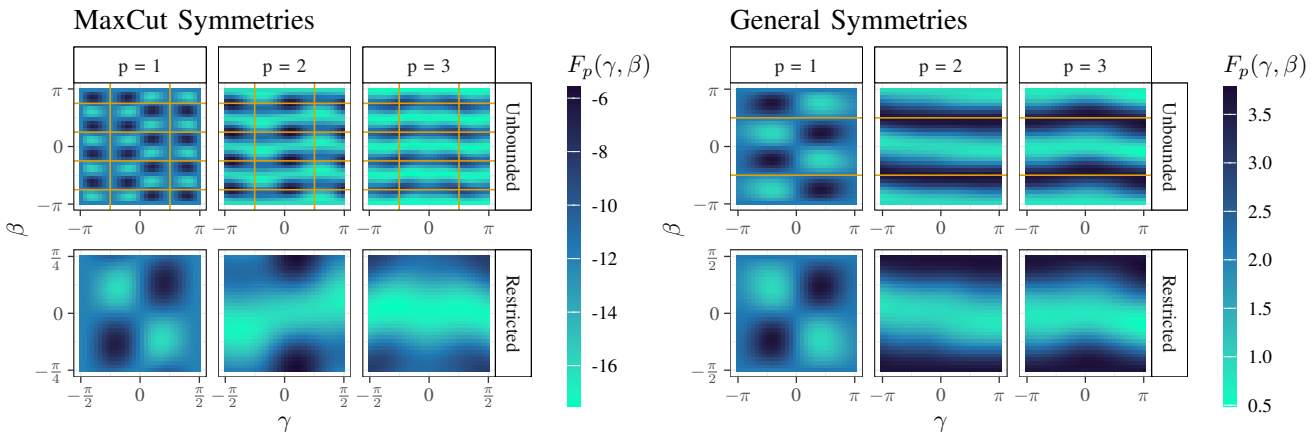

FIGURE 2: Symmetries specific to MaxCut (lhs) and general QAOA (rhs) illustrated by a single 16-vertex, 3-regular instance of MaxCut and a single 16 qubit Max3SAT instance with $\alpha = 4.\overline{3}$. *Top rows:* Overall optimisation landscape $F_p(\gamma, \beta)$ with $\gamma, \beta \in [-\pi, \pi]$ and symmetry axes marked in orange. *Bottom rows:* Restriction to the subset $\gamma \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\beta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ for MaxCut and $\gamma \in [-\pi, \pi]$, $\beta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ in general that captures all information. The landscapes are obtained by the sequential method.

and two vertical ($[-\frac{\pi}{2}, \frac{\pi}{2}]$ range vs. $[\frac{\pi}{2}, \pi]$ and $[-\pi, -\frac{\pi}{2}]$ ranges) identical sections, respectively. If we account for the symmetries and eliminate the corresponding degeneracies, the landscape can be restricted to the subset shown in the lower part of of the figure.

Additionally, QAOA parameters are invariant under time reversal, which is visualised by the fact that in the $p = 1$ landscapes, quadrants $(-\beta, -\gamma)$ and (β, γ), as well as quadrants$(-\beta, \gamma)$ and $(\beta, -\gamma)$ are point symmetric. Notable, for $p > 1$, time reversal symmetry may not always be visible in the plots, since we are only viewing a small part of the entire optimisation landscape given by all components of $\vec{\gamma}, \vec{\beta}$.

B. MAXCUT

Figures 3 and 4 show average residual energy $\bar{r}$ and standard deviation σ_r for our MaxCut instances. Each plot facet corresponds to the optimisation landscape of $\vec{\beta}^{(p)}, \vec{\gamma}^{(p)}$ with $\beta_i^{(p)}, \gamma_i^{(p)} (0 \leq i < p)$ components fixed to parameters of interest. Columns vary depth p, and different rows correspond to different parameter initialisation methods. The average quality of the results produced by the parameters fixed at each layer by these methods is shown in Fig. 5. For comparison, the results for one individual instance are shown in the appendix in Sec. C-A.

a: Average Residual Energy Landscapes

The parameter landscapes obtained with the sequential method, COBYLA, and $\text{LR}_{-\beta}$ share common features: the optimal absolute value of β decreases with increasing p, while the optimal absolute value of γ, which grows at shallow depth, becomes effectively arbitrary at larger p once $\beta \approx 0$ (see Fig. 3). In other words, the average residual energy landscapes given by the individual components of β, γ become increasingly invariant under γ as β approaches zero. This suggests that as β approaches zero, the system state is left unchanged by further temporal propagation under the cost Hamiltonian, while propagation under the mixer Hamiltonian still induces dynamics. Intuitively, if the system sufficiently approximates an energy eigenstate of the cost Hamiltonian, then adding an additional layer of cost and mixer operators leaves the state invariant under the cost Hamiltonian. In this regime, all choices of γ are therefore equally effective. Conversely, further propagation under the mixer Hamiltonian will change the state, so parameters β need to be close (or equal) to zero for the best results. Accordingly, if the depth is sufficient for a QAOA circuit with a given set of parameters to adequately approximate one of the eigenstates of H_C, we can expect the landscape of the highest depth components of γ, β to converge to such a landscape. In contrast, the landscapes around the $\text{LR}_{+\beta}$ parameters show the opposite of these convergence properties, with a choice of a small absolute value for β leading to a *high* $\bar{r}$, and the optimal choice of β tends towards high values with increasing p. This suggests that, rather than converging towards the ground state, the system approaches an excited state, which is natural, since a positive $\Delta\beta$ corresponds to a positive mixer with $|+\rangle$ as its maximally excited state. and with $\Delta\gamma > 0$, this corresponds to maximising H_C.

The sequential method is comparatively the fastest to converge towards a landscape, where the energy is invariant under γ, which reaches such a state at a depth of $p = 4$. Fixing β at (or very close to) zero (cf. Fig. 15) at higher depths $p \geq 4$ results in neither an increase nor decrease in $\bar{r}$ and no noticeable change in the landscapes. When optimising the parameters using COBYLA, this transition between landscapes takes longer, reaching a landscape

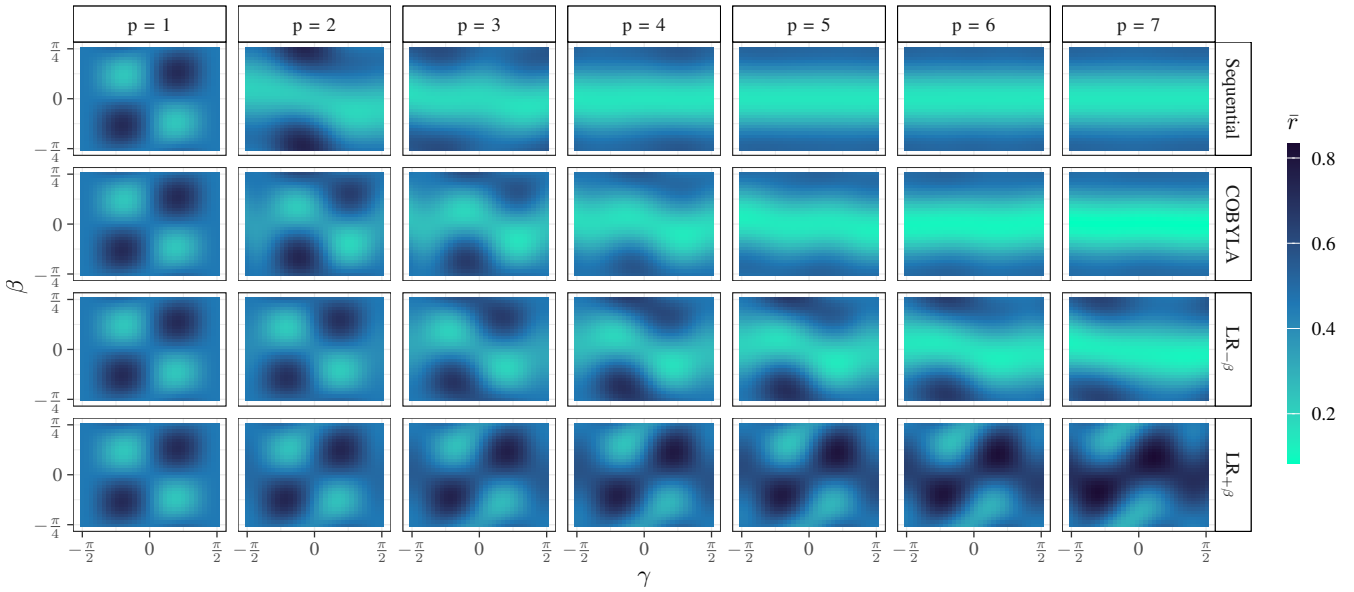

FIGURE 3: Average residual energy $\bar{r}$ for 40 3-regular MaxCut instances of sizes 10 to 16, with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

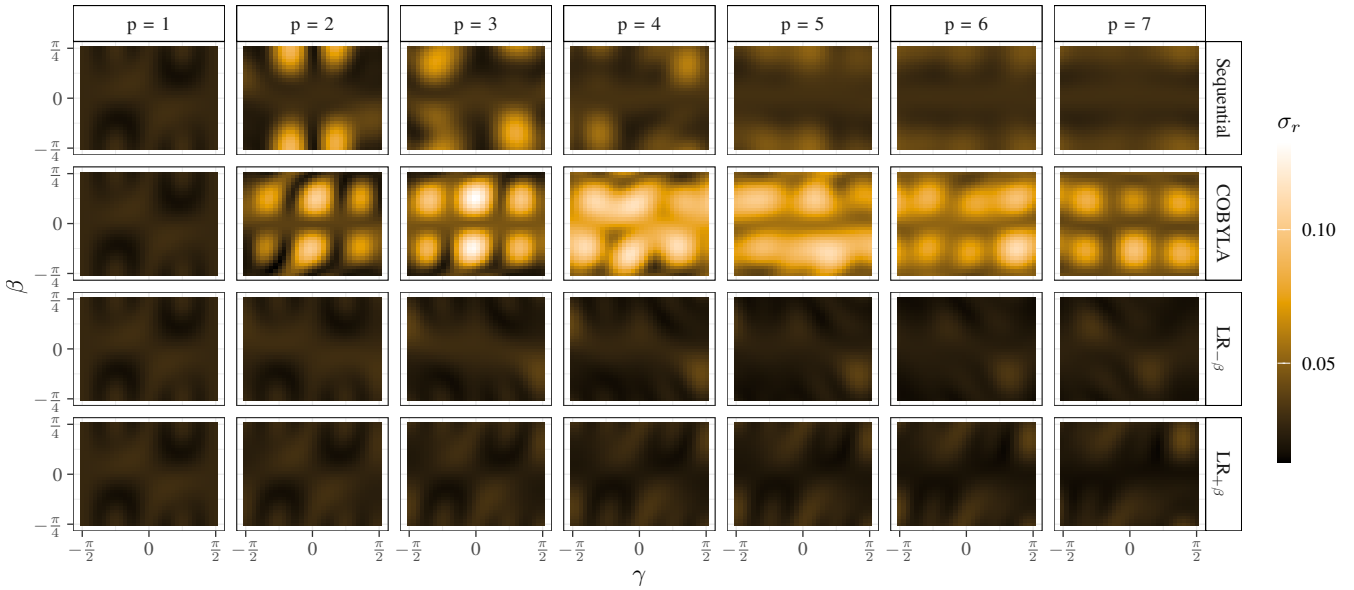

FIGURE 4: Standard deviation of the residual energy $\bar{r}$ in Fig. 3 over 40 MaxCut instances.

similar to the $p = 4$ sequential landscape at $p = 7$, with more intermediate transition steps. The resulting $\bar{r}$ is higher than that of the sequential method at low depth ($p < 4$), but starts outperforming the sequential method as the depth increases ($p > 4$). Unlike the sequential method, the quality keeps improving throughout all layers up to $p = 7$ when optimising (see Fig. 5). The LR$_{-\beta}$ parameters take the longest to converge to a state that is invariant under the phase operator, and in fact do not quite reach this point within $p = 7$ with the given $\Delta\beta, \Delta\gamma$. As shown in Fig. 5, the average residual energy consistently improves at a similar rate to the optimised parameters, with the optimised parameters performing slightly better at all depths. At $p \geq 6$, The LR$_{-\beta}$ parameters outperform the sequential parameters. The average residual energy given by the LR$_{+\beta}$ parameters changes inversely to the LR$_{-\beta}$, albeit the landscapes of these methods are not an inverse of each other, with the individual LR$_{+\beta}$ landscapes differing less from preceding ones than for the LR$_{-\beta}$ parameters.

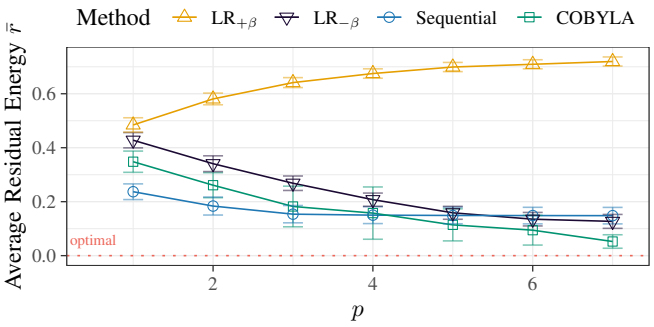

FIGURE 5: Average and standard deviation of the approximation quality for 40 3-regular MaxCut instances of sizes 10 to 16 when fixing parameters at each layer p according to the considered methods.

b: Convergence to γ Invariance at $\beta = 0$

Another notable observation from Fig. 3 is that the maxima seem to shift in accordance with the changes in γ and β with increasing p: (1) the maxima shift towards the top/bottom of the landscape (*i.e.*, the symmetry boundary of β) and gradually diminish, while minima shift towards $\beta = 0$; and (2) the maximal/minimal regions, which are duplicated due to time-reversal symmetry, broaden diagonally and merge into a single maximal/minimal region.

For LR$_{-\beta}$, higher depths correspond to a high γ relative to β and lower depths to a low γ relative to β. Comparing theses ranges illustrates that the minima/maxima shifts are consistent with the changes in β, γ: At low depth ($p < 4$), where γ is (near) zero, property (2) is more pronounced than property (1), with the position of the extrema barely changing. Only the width of the minimal/maximal region increases along the diagonal between duplicate minima/maxima. At higher depths ($p > 4$), as γ increases (and β approaches zero), property (1) becomes more pronounced. The sequential method fixes non-zero values for both β and γ at $p = 1$ and subsequently, at $p = 2$, reaches a landscape similar to the $p = 6$ LR$_{-\beta}$ parameters showing both properties (1) and (2). The lower depth ($p < 4$) scans show that the COBYLA optimiser has a tendency to chose higher values of β and lower values of γ than the sequential method (cf. Fig. 15). Furthermore, property (2) is more pronounced for the sequential parameters than those obtained with COBYLA, which at low depth have a longer propagation under β relative to the chosen γ. This indicates that for COBYLA property (1) becomes more pronounced as β decreases, relative to γ.

c: Residual Energy Standard Deviation Landscapes

At $p = 1$, the standard deviation (see Fig. 4) from $\bar{r}$ is low ($\sigma_r \leq 0.05$) across the entire landscape, with areas of almost zero σ_r in parts of the landscape where $\beta \approx \pm\frac{\pi}{8}, \gamma \approx \pm\frac{\pi}{4}$. The parts of the landscape which correspond to the minima in the $p = 1$ $\bar{r}$ landscape in Fig. 3 result in σ_r not significantly different from the rest of the landscape (≤ 0.05). At $p = 2$ to $p = 4$, for the sequential method, there are some areas of moderate standard deviation ($\sigma_r \leq 0.1$) that correspond to a long propagation under $\beta \approx \pm\frac{\pi}{4}$ and specific values of γ. In the corresponding $\bar{r}$ landscapes of Fig. 3, these regions of moderate σ_r contain maxima or medium $\bar{r}$. At $p > 4$, σ_r is more uniform across the entire landscape, with the overall magnitude σ_r decreasing with increasing p, but with higher σ_r as in the $p = 1$ landscape. When optimising the sequential parameters using COBYLA, σ_r is comparatively higher ($0.1 \leq \sigma_r \leq 0.15$ in some areas of the landscapes between $p = 3$ and $p = 6$) and less uniformly distributed. There are some narrow regions where σ_r is lower than the rest of the landscape, namely in the areas where $\beta = 0$ or $\beta = \pm\frac{\pi}{4}$ and in areas where $\gamma \approx \pm\frac{\pi}{4}$ (the exact γ seems to vary slightly), making the landscape appear segmented into areas of high σ_r. Both LR parameters result in landscapes with low, uniform σ_r that barely differ from the $p = 1$ landscape, with σ_r only slightly decreasing in some areas and slightly increasing in others. For the LR$_{-\beta}$ parameters, the areas in the middle of the landscape increase in σ_r, while for the LR$_{+\beta}$ parameters the σ_r decreases in these areas. For the areas at the edges of the plot, the inverse of this applies.

C. MAXCUT: HIGHER DEPTHS

To examine the parameter landscapes at higher depths we further evaluate the same MaxCut instances up to $p_{\text{target}} = 21$. The results of these experiments are shown in Figures 6 and 7, respectively corresponding to the average residual energy $\bar{r}$ and the standard deviation σ_r. The landscape is evaluated in intervals of two layers for COBYLA and the LR parameters, omitting every second landscape, to save computation time due to the increased number of circuit evaluations and circuit depths. The average quality of the energies produced by the parameters fixed at each layer by these methods is shown in Fig. 8. The results for one individual instance are shown in Sec. C-A.

a: Average Residual Energy Landscapes

At higher depths, the landscape around the sequential parameters (top row of Fig. 6) does not change in a noticeable way compared to the $p_{\text{target}} = 7$ landscape, with the lowest energy values occurring at $\beta = 0$ and arbitrary γ. As with the lower depth experiments (Sec. V-B), from $p = 4$ onwards, the sequential method consistently fixes $\beta = 0$ (cf. Fig. 15), that is, no further temporal propagation under the mixer occurs at higher depths, and since any propagation under the phase operator leaves the system in its state, the residual energy also remains constant (cf. Fig. 8). This behaviour aligns with our expectations, since the sequential method always chooses the best available parameters at each p without modifying previously fixed parameters. So, once it has reached a landscape where further propagation under β only results in higher energies and the choice of γ is arbitrary, there is no further propagation under the mixer at higher depths.

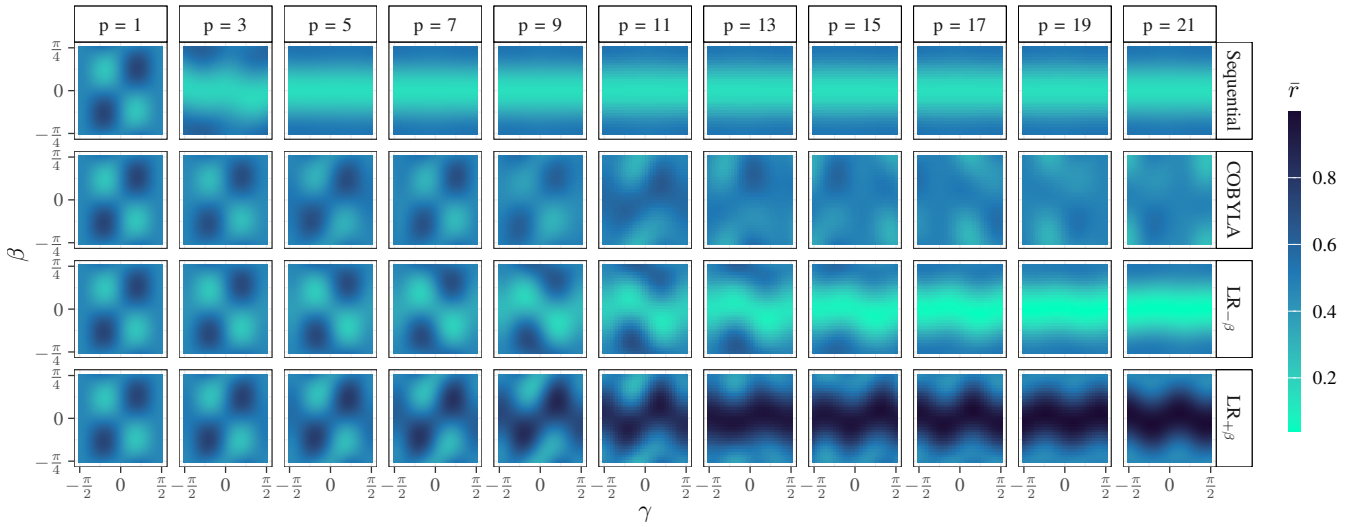

FIGURE 6: Average residual energy $\bar{r}$ for the same 40 MaxCut instances of sizes 10 to 16 shown in Fig. 3 with $p_{\text{target}} = 21$, with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

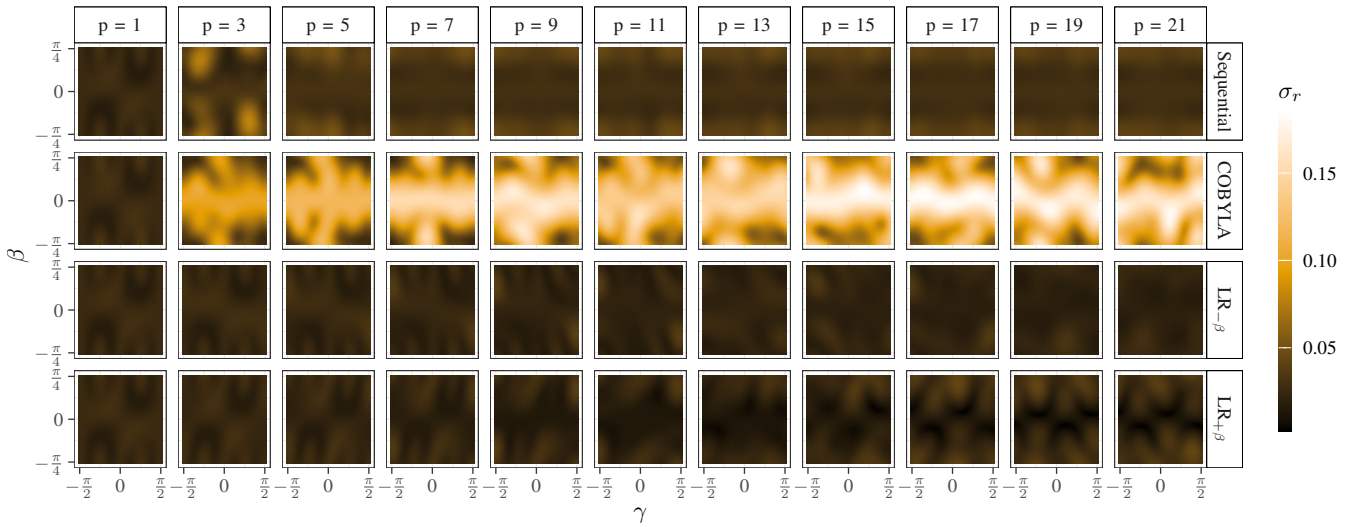

FIGURE 7: Standard deviation of the residual energy $\bar{r}$ in Fig. 6 over 40 MaxCut instances, where $p_{\text{target}} = 21$.

The landscapes around the LR$_{-\beta}$ parameters (lower-middle row of Fig. 6) progress similarly to the landscapes in the lower depth experiments, but at an overall slower rate. This is natural, as the slope of the linear ramp is $\frac{\Delta}{p}$, and using a higher p with the same Δ corresponds to a more gradual slope, which in turn corresponds to more gradual changes in the variational parameters. The higher depth landscape at $p > 7$ continues to converge to a landscape that is dependant on β and invariant under γ, and unlike the lower depth case, there is no more variation in $\bar{r}$ depending on γ at $p = 21$. The performance of the higher depth LR$_{-\beta}$ parameters (see Fig. 8) at low depth $p < 13$ is worse than the low depth LR$_{-\beta}$ parameters (compare Fig. 5), but after $p = 13$, the higher depth LR$_{-\beta}$ parameters start outperforming the lower depth variant, as well as the sequential parameters. At $p \geq 17$ a lower $\bar{r}$ can be achieved than the best-performing optimised parameters at lower depth (cf. Fig. 5).

The LR$_{+\beta}$ landscape (bottom row of Fig. 6) similarly progresses at a slower rate compared to to the respective lower depth landscapes. The final landscape is invariant under γ close to $\beta = 0$, albeit there is some variation in $\bar{r}$ close to $\beta = \pm\frac{\pi}{4}$ depending on γ.

When using COBYLA to optimise the LR$_{+\beta}$ parameters, the resulting landscapes are distinct from the case, where we initialised the optimiser with sequential parameters at $p_{\text{target}} = 7$. The landscapes mainly consist of rather average $\bar{r}$ and the maxima and minima of the individual landscapes

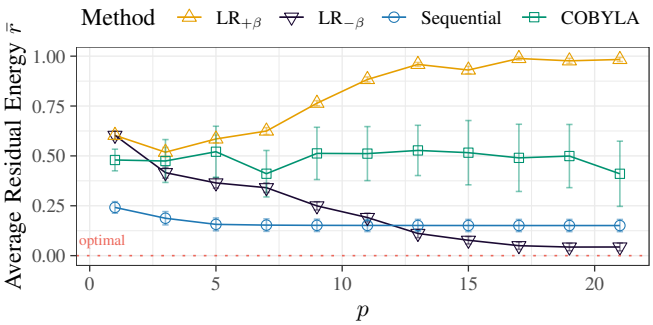

FIGURE 8: Average and standard deviation of the approximation quality for 40 3-regular MaxCut instances of sizes 10 to 16 when fixing parameters at each layer p according to the considered methods, where $p_{\text{target}} = 21$.

are less pronounced, with convergence patterns becoming less discernable as p increases. At $p \leq 11$, the landscapes progress similarly to the LR$_{+\beta}$ landscapes. However at $p > 11$, the maxima become less pronounced, compared to the the LR$_{+\beta}$ parameters, and the minima shift horizontally towards the corners of the landscape at $p = 21$, which corresponds to a long propagation under both β and γ. Overall, the optimiser performs worse than the sequential parameters and the LR$_{-\beta}$ parameters (see Fig. 8), as well as the optimised sequential parameters in the lower depth case (cf. Fig. 5).

b: Residual Energy Standard Deviation Landscapes

The standard deviation of these landscapes (see Fig. 7) at higher depths behaves similarly to the $p_{\text{target}} = 7$ experiments: The sequential parameters and the LR$_{-\beta}$ parameters both result in a low standard deviation across the entire landscape, with $\sigma_r < 0.1$ and $\sigma_r \leq 0.05$, respectively. LR$_{+\beta}$ parameters also result in a similar σ_r as the lower depth case (cf. Fig. 4), but with slightly higher $\sigma_r < 0.1$ occurring in the landscape where β is non-zero. The landscapes of COBYLA initialised with LR$_{+\beta}$ parameters have the highest $\sigma_r < 0.2$, across a large part of the landscape, with the highest σ_r generally occurring in the part of the landscape that corresponds to a short propagation under the mixer, that is, where β is close to zero. This indicates that there is a relatively high variation in the cost landscape across MaxCut instances when optimising LR$_{+\beta}$ parameters. Notably, there are no areas of low σ_r at specific values of γ or β, such that the σ_r landscape is divided into segments, as was observed in the lower depth experiments, where COBYLA was initialised with the sequential parameters (see Sec. V-B).

D. VERTEXCOVER

Figures 9 and 10 respectively show the average residual energy $\bar{r}$ and the standard deviation σ_r of the VertexCover instances on the 40 3-regular graph instances used in Sections V-B and V-C, up to $p_{\text{target}} = 7$. The average quality of the energies produced by the parameters using the described methods is shown in Fig. 11. The results for a single instance are shown in Sec. C-B. Aside from the fact that here we initialise COBYLA with LR$_{-\beta}$ parameters, the figures use the same methods and follow the same structures as the MaxCut experiments (cf. Sec. V-B).

a: Average Residual Energy Landscapes

In comparison to MaxCut, the $p = 1$ landscape for VertexCover exhibits a higher density of local minima and maxima. A possible reason for this is that there are more computational basis states corresponding to invalid solutions for VertexCover than for MaxCut. A slight variation to the parameters may cause a transition from a valid solution state to an invalid solution state, leading to irregularities in the surface of the optimisation landscape. Furthermore, as discussed in Sec. V-A, the range in which γ, β are periodic is different to MaxCut, as only general symmetries apply for VertexCover. Nevertheless, the $p = 1$ VertexCover landscape shares some general properties with the $p = 1$ MaxCut landscape, such as time-reversal symmetry. Additionally, the landscape follows the general structure, with the lower left and upper right quadrants in general containing lower values of $\bar{r}$ than the lower right and upper left quadrants.

As p increases, the sequential method quickly converges to a landscape with a smoother surface, where $\bar{r}$ is invariant under γ when $\beta = 0$ or $\beta = \pm\frac{\pi}{2}$, at $p \geq 2$. However, unlike in the MaxCut experiments, there is no examined depth, where the landscape is entirely invariant under γ for any choice of β. From $p \geq 3$ onwards, the sequential method fixes β close to zero (cf. Fig. 15) and $\bar{r}$ remains constant (cf. Fig. 11), but the location and magnitude of the maxima at the edges of the landscape still change depending on the chosen γ. This suggests that the system is in a state, in which further propagation under the phase operator will leave the system in the state, if there is no propagation under the mixer. Evolution under the mixer in turn may change the state depending on the magnitude of both γ and β, indicating that the propagated state no longer closely approximates an eigenstate of H_C. In short, the effect of propagation under β is more pronounced than for the MaxCut experiments (see Sec. V-B and Sec. V-C).

In the LR$_{-\beta}$ landscape (third row of Fig. 9) at $p \geq 2$ the minima also shift towards the center of the plot, but in comparison to the landscape of the sequential method, it is overall less smooth and the minima and maxima are less extensive. Furthermore, while the parts of the landscape where $\beta = 0$ contain similarly low values of $\bar{r}$, the landscape is never completely invariant under γ in these regions, suggesting that the LR$_{-\beta}$ fail to converge to a state which closely approximates an eigenstate of H_C. This landscape barely changes at higher depths $p \geq 3$.

When optimising the LR$_{-\beta}$ parameters using COBYLA (upper-middle row of Fig. 9), the $p = 2$ and $p = 3$ landscapes are similar to the LR$_{-\beta}$ landscapes for the same p. However, as p increases further, the landscapes converge to a state that is invariant under γ at $\beta = 0$. Compared

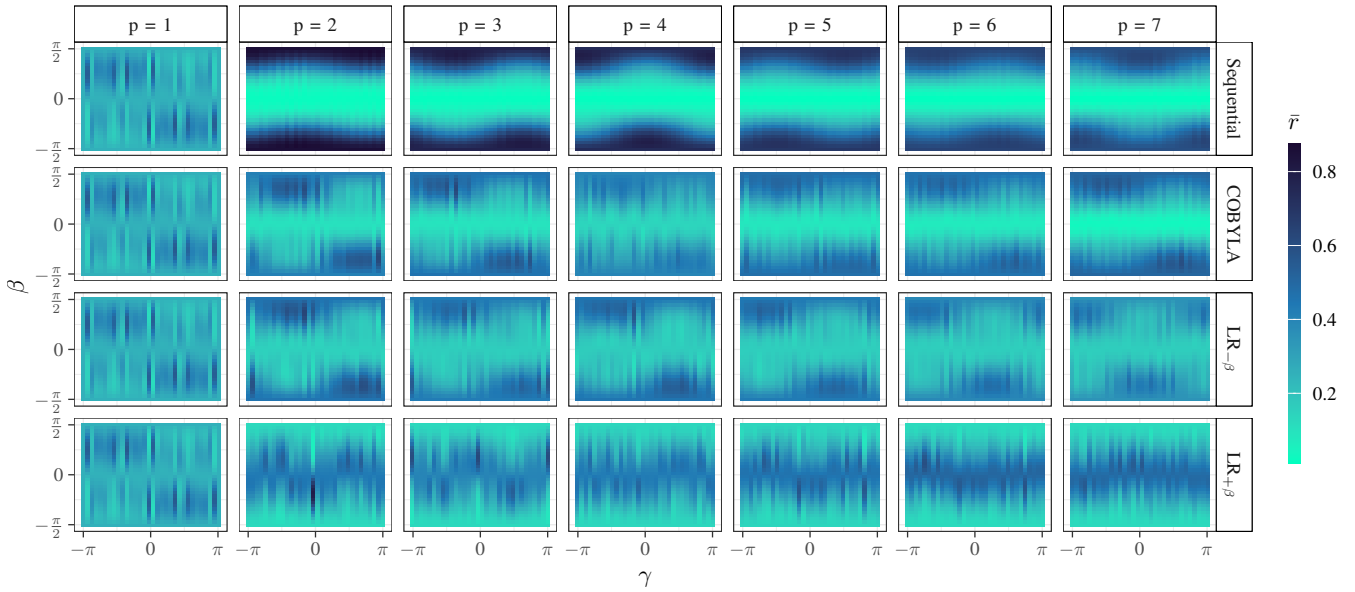

FIGURE 9: Average residual energy $\bar{r}$ for 40 VertexCover instances of sizes 10 to 16, with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

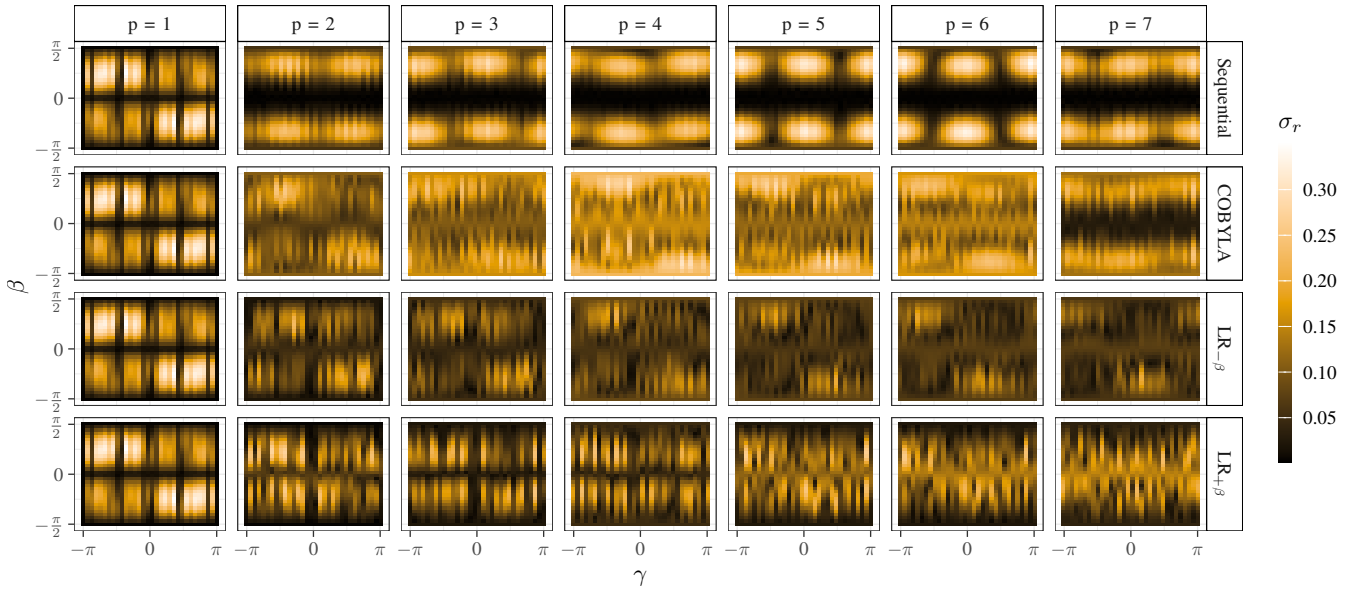

FIGURE 10: Standard deviation of the residual energy $\bar{r}$ in Fig. 9 over 40 VertexCover instances.

to the landscapes of the sequential method, the optimiser landscapes are less smooth and the minimum is narrower in β, with less pronounced maxima up to $p = 7$. Fig. 11 also shows a similar performance of the COBYLA and sequential method at $p = 7$.

The LR$_{+\beta}$ parameters (bottom row of Fig. 9), converge towards the inverse of the other landscapes, and similar to the LR$_{-\beta}$, do not reach a landscape which is invariant under γ when $\beta = 0$, instead only reaching a landscape with low variance under γ at $\beta = 0$. Unlike the LR$_{-\beta}$ parameters, the $\bar{r}$ of the LR$_{+\beta}$ keeps changing with increasing p, towards higher values of $\bar{r}$ (see Fig. 11). Notably, at $p = 2$ with γ slightly below zero, there is a relatively high maximum for negative β and a relatively low minimum for positive β. These extrema are mirrored in β at $p = 3$, suggesting that some specific combinations of γ, β at varying depth lead to opposite behaviours.

b: Convergence to γ Invariance at $\beta = 0$

As with MaxCut, the maxima and minima in the landscapes seem to shift and broaden depending on the magnitude of γ, β, although this is less evident due to the dense

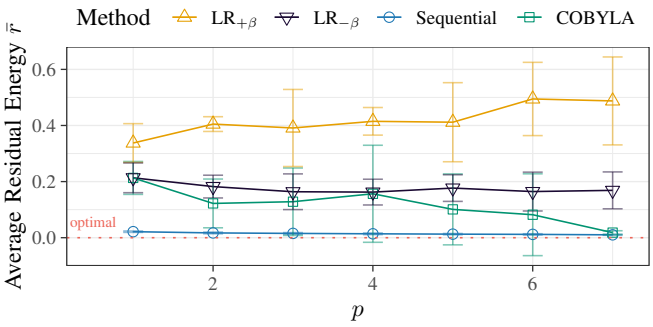

FIGURE 11: Average and standard deviation of the approximation quality for 40 VertexCover instances of sizes 10 to 16 when fixing parameters at each layer p according to the considered methods.

local minima/maxima in the landscape. The most notable example here is evident in the sequential method, for which β quickly decreases to zero within $p \leq 3$ (cf. Fig. 15), as the maximal/minimal regions quickly merge accordingly. As the state is further propagated under γ, the maxima become diminished and tend towards the vertical extremes of the landscape while the minima become narrower around $\beta = 0$.

c: Residual Energy Standard Deviation Landscapes

The standard deviation landscape (cf. Fig. 10) at $p = 1$ shows a relatively high σ_r in the upper left and lower right quadrants, which tend to contain higher values of $\bar{r}$ in the average landscapes (cf. Fig. 9). In the other quadrants, which tend to contain lower $\bar{r}$, σ_r is medium in magnitude. In the areas, in which $\beta = 0$ and $\beta = \pm\frac{\pi}{2}$ with medium $\bar{r}$, or in areas with pronounced maxima and minima of $\bar{r}$, σ_r is close to zero. At higher depths, for the sequential method σ_r is close to zero in the regions where β is (near) zero, and $\bar{r}$ is minimal. The standard deviation is higher and more variable in the areas where β is closer to $\pm\frac{\pi}{2}$ and $\bar{r}$ is higher. Notably, the landscapes, which the sequential method converges against, lie in areas of low σ_r close to $\gamma = \pm\frac{\pi}{2}$ for $p = 5$ and $p = 6$, though this effect is less recognizable at $p = 7$. The LR parameters result in landscapes where middling and higher values of σ_r are distributed more uniformly across the landscape, with frequent narrow areas of low standard deviation occurring in irregular intervals of γ. In general, for the LR$_{-\beta}$ parameters, higher σ_r lie in areas corresponding to a high $\bar{r}$, and more middling σ_r in areas with low $\bar{r}$. The areas with low σ_r are rather uneven, partly corresponding to more pronounced extrema and partly to a medium $\bar{r}$. The LR$_{+\beta}$ parameters result in landscapes where areas of low σ_r seem to appear more regularly at low p near $\beta = 0$ and in areas with pronounced extrema of $\bar{r}$, as well as at high p, in areas of the where $\beta = \pm\frac{\pi}{2}$. However, overall high values of σ_r tend to occur more frequently and to a greater extent than in the LR$_{-\beta}$ landscapes. When optimising the LR$_{-\beta}$ parameters with COBYLA, the standard deviation of the resulting landscapes increases noticeably and is spread more uniformly across the entire landscape as p increases. Notably, after fixing $p = 6$ optimised parameters, a transition occurs in the σ_r landscape, causing the $p = 7$ landscape to be more similar to that of the sequential parameters at $p = 7$, with low σ_r in the region where β is near 0, and higher σ_r where β is closer to $\pm\frac{\pi}{2}$.

E. MAX3SAT

Figures 12 and 13 show the average residual energy $\bar{r}$ and the standard deviation for 10 "hard" Max3SAT instances with $\alpha \in (3.5; 4.9]$ [31], up to $p_{\text{target}} = 7$. The average quality of the energies produced by the parameters fixed at each layer is shown in Fig. 11. The results for a single instance are shown in Sec. C-C. Apart from initialising the COBYLA optimiser with LR$_{+\beta}$ parameters, the figures use the same methods and follow the same structures as the previous experiments (cf. Sec. V-B or Sec. V-D). Additionally, we evaluated 10 Max3SAT instances from the easy range, $\alpha \notin (3.5; 4.9]$, which gave qualitatively similar results, as we describe in Appendix B.

a: Average Residual Energy Landscapes

The resulting landscapes resemble those of MaxCut, but exhibit key differences: As discussed in Sec. V-A, the MaxCut specific symmetries do not apply for Max3SAT; only the general β-symmetry and time reversal symmetry applies. Compared to MaxCut, the Max3SAT landscapes contain fewer regions of intermediate $\bar{r}$, dominated instead by high and low values. At higher p, the maxima are larger, whereas for MaxCut the highest $\bar{r}$ decrease with depth (cf. Sec. V-B). Overall, all methods perform worse than in previous experiments, generally yielding higher (or lower for LR$_{+\beta}$) $\bar{r}$ (cf. Fig. 14 and Fig. 5, also cf. Fig. 8 and Fig. 11). This suggests that for Max3SAT, inadequate temporal propagation (*i.e.*, the system evolves away from the target state) at higher depth result in higher $\bar{r}$ than for inadequately propagated QAOA in MaxCut and Vertexcover at the same depths. The sequential method landscape (top row of Fig. 25) progresses similarly to prior experiments, approaching a γ-invariant and β-periodic structure, with minima at $\beta = 0$ and maxima at $\beta = \pm\frac{\pi}{2}$. Unlike other cases, the sequential method's average $\bar{r}$ continues to decrease up to $p = 7$, though the improvement becomes negligible beyond $p = 5$ (cf. Fig. 14). Analogously, the LR landscapes evolve more slowly, showing intermediate $\bar{r}$ as p increases. At $p = 7$, neither the LR$_{-\beta}$ nor the LR$_{+\beta}$ have converged to γ-invariance at $\beta = 0$: For LR$_{-\beta}$, the best parameters occur near $\beta \approx \frac{\pi}{8}$, $\gamma \approx -\pi$ or $\beta \approx -\frac{\pi}{8}$, $\gamma \approx \frac{\pi}{3}$; for LR$_{+\beta}$ the extrema appear at the corresponding inverted-sign regions. The COBYLA optimiser, initialised with the LR$_{+\beta}$ parameters, produces increasingly, but intermediate $\bar{r}$, and weaker extrema as p increases.

None of the methods closely approximates the true ground state energy (cf. Fig. 14). The sequential method performs best, saturating beyond $p = 5$; LR$_{-\beta}$ ranks second

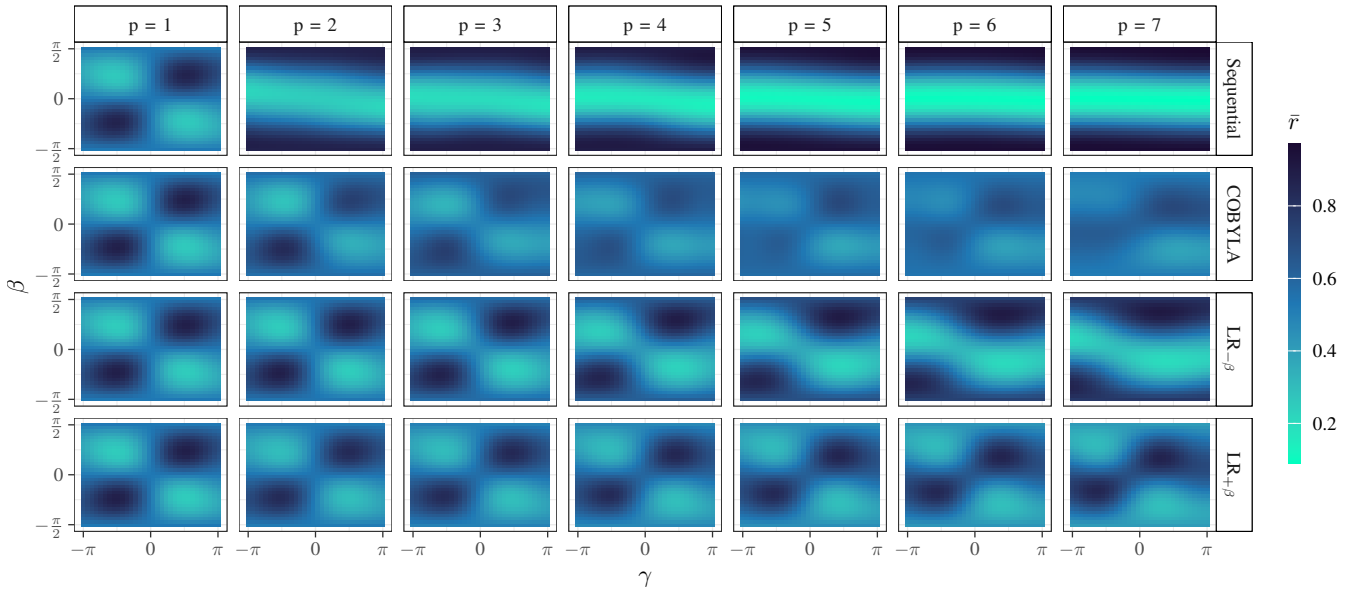

FIGURE 12: Average residual energy $\bar{r}$ for 10 Max3SAT instances of sizes 14 to 24, with $\alpha \in (3.5; 4.9]$ and γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

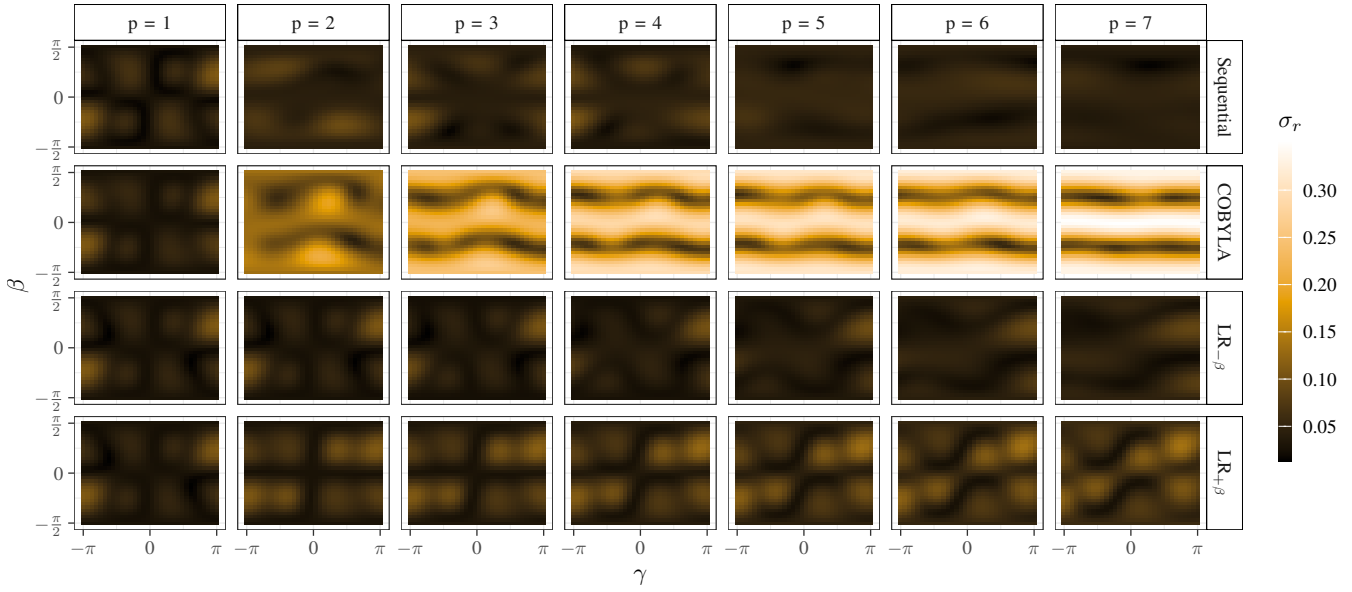

FIGURE 13: Standard deviation of the residual energy $\bar{r}$ in Fig. 12 over 10 Max3SAT instances.

for $p < 7$, improving comparatively to the sequential method up to $p = 4$ but starting from a worse $\bar{r}$ at $p = 1$. COBYLA, initialised with energy maximising LR$_{+\beta}$ parameters, performs similarly to LR$_{-\beta}$ at $p \leq 2$, worse from $p = 3$ to $p = 6$, and substantially better at $p = 7$, surpassing LR$_{-\beta}$, but performing worse than the sequential method.

b: Convergence to γ Invariance at $\beta = 0$

As in Sec. V-B, V-C and V-D, landscape extrema shift and broaden with increasing γ and decreasing β. The LR parameters illustrate this gradual transition: At $p = 2$, for high β and small γ, the landscape resembles that of $p = 1$. With increasing p, duplicate extrema broaden diagonally and merge. Whether minima or maxima merge across the central diagonal depends on the β ramp sign: minima merge for negative β, maxima for positive β. For $p \geq 4$, as $\beta \to 0$ and γ increases, extrema shift vertically, with minima moving

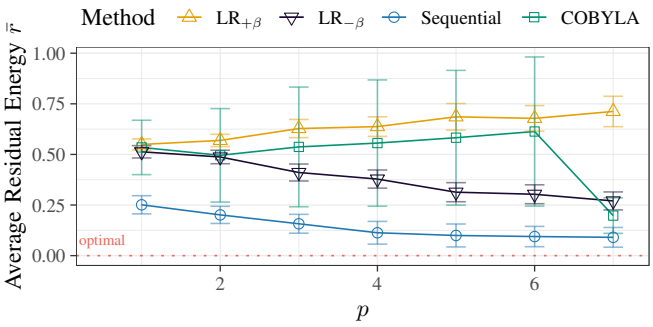

FIGURE 14: Average and standard deviation of the approximation quality for 10 Max3SAT instances of sizes 14 to 24 with $\alpha \in (3.5; 4.9]$ when fixing parameters at each layer p according to the considered methods.

inward and maxima outward for negative β, with the reverse for positive β. These shifts intensify as γ approaches Δ_γ.

c: Residual Energy Standard Deviation Landscapes

At $p = 1$, the standard deviation (cf. Fig. 13) is generally relative low, but slightly higher than for $p = 1$ MaxCut (cf. Sec. V-B,V-C) and lower than for $p = 1$ VertexCover (cf. V-D). The largest σ_r appear near $\gamma = \pm\pi$ and $\beta = \pm\frac{\pi}{4}$, coinciding with the extrema of the $p = 1$ $\bar{r}$ landscape (cf. Fig. 12). Regions near $\beta = 0$ or $\gamma = 0$ show nearly zero σ_r, segmenting the landscape into high- and low-variance zones. As p increases, the sequential method's landscape becomes more uniform: peak σ_r remains stable up to $p = 4$ and decreases slightly from $p \geq 5$. The LR landscapes, homogenise more slowly; the LR$_{-\beta}$ landscapes change minimally up to $p = 3$, while LR$_{+\beta}$ landscapes transition slightly faster at $p = 2$, but then plateau. Overall, LR$_{+\beta}$ exhibits higher σ_r than LR$_{+\beta}$. For COBYLA initialised with LR$_{+\beta}$ parameters, higher-depth ($p \geq 2$) landscapes contain high σ_r, approaching ≈ 0.35 as p increases. Low-variance regions ($\sigma_r \leq 0.1$) near $\beta = \pm\frac{\pi}{4}$ become progressively less dependant on γ, so that by $p = 7$ σ_r varies little with γ. This suggests that higher-depth runs with $\beta = \pm\frac{\pi}{4}$ yield r-values closely clustered around $\bar{r}$. However, this does not necessarily indicate that an eigenstate of H_C is being approximated as $\bar{r}$ is still γ-dependant.

F. (NON-)PATTERNS IN QAOA PARAMETERS

To assess how well the examined parameters conform to the patterns described in the literature, Fig. 15 shows the parameters determined by the sequential method (left) and COBYLA (right). The LR ramps are omitted, as they follow a linear pattern by construction.

For the sequential method, the β parameters follow the expected pattern, decreasing (non-linearly), rapidly at low p and gradually at higher p. The γ parameters adhere to literature trends only for low depth ($p \leq 3$ for MaxCut, $p \leq 2$ for Max3SAT). At higher depth and for $p \geq 1$ VertexCover, γ varies irregularly, often fixed to $\pm\pi$ for larger $p \geq 2$ Max3SAT instances. These deviations likely arise because the sequential method quickly converges to a γ-invariant state when $\beta = 0$ (cf. Sec. V-B, V-C, V-D and V-E), rendering γ arbitrary at high p. For VertexCover, inconsistencies at low depth stem from $\beta = 0$ being fixed from $p = 2$ onward; since the landscape is then γ-invariant (cf. Sec. V-D), γ becomes arbitrary and no clear parameter pattern emerges.

When COBYLA is initialised with the sequential parameters (top row), the resulting optimised parameters show similar behaviour, deviating only moderately, but with less smooth changes in β, γ. Initialisation with LR$_{-\beta}$ (third row) leads to greater deviations, generally breaking the patterns in both β and γ, except in 10 qubit instances where patterns persist in β and in γ up to $p \leq 3$. Initialisation with LR$_{+\beta}$ (second and bottom rows), which corresponds to maximising H_C and thus a disadvantageous starting point, produces deviations larger than with the sequential initialisation. The largest discrepancies occur in the first and last β, γ components. For Max3SAT, optimisation from the LR$_{+\beta}$ parameters yields slightly greater deviations for smaller instances than in than for $p_{\text{target}} = 21$ MaxCut, though less than when initialised with LR$_{-\beta}$. Since these optimised parameters generally yield intermediate average residual energy (cf. Sec. V-C and V-E), and vary little for larger instances, the optimiser is more likely to getting stuck in a local optimum of the *entire* optimisation landscape, that is, the landscape given by all p components of β, γ. Combined with the observation that large β and γ values are typically fixed for the final layer (cf. Sec. V-C, V-E), this suggests that the optimiser converges to one of the minima located at the edges of the final landscapes.

VI. DISCUSSION

Our results paint a nuanced picture about advantages and disadvantages of QAOA variants: Despite known uniform, generic and instance-independent characteristics of QAOA at unit depth $p = 1$ [25, 40], we find higher circuit depths exhibit an influence of both, the combinatorial *problem*, and the specific *instance*. Nevertheless, commonalities are still observable for deeper circuits, particularly when fixing lower-depth parameters to certain initial values.

Most notably, parameters resulting in low residual energies at every depth lead to a converging landscape that is invariant under γ and periodic in β with a minimum at $\beta = 0$ and maxima towards the top and bottom of the symmetry range. As p increases, the choice of γ becomes effectively arbitrary, yielding equivalent energies across all values at large depth. Conversely, the optimal choice of β approaches zero as p increases. The number of steps required for the landscapes to converge to such a state depends on the relative magnitudes of β, γ: For instance, linear ramps with decreasing β and increasing γ starting from 0 converge slower than methods starting with a non-zero $\gamma \geq \beta$ (*e.g.* the sequential method, or COBYLA using favourable initial parameters). These "faster" methods

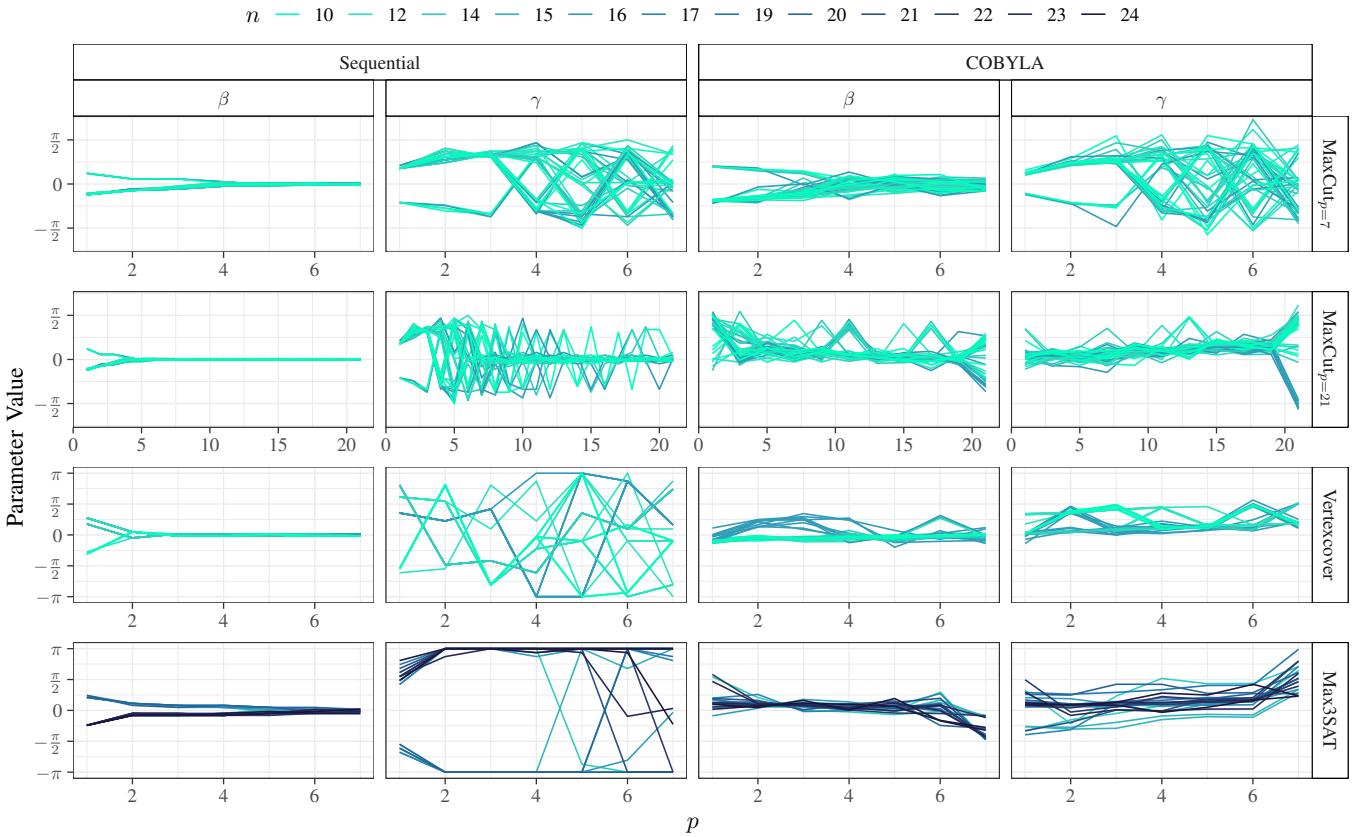

FIGURE 15: Parameter components of $\vec{\gamma}, \vec{\beta}$ at each p determined by our sequential method and the COBYLA optimiser for each instance of the considered optimisation problems. Colour denotes number of qubits n for each problem instance. For each problem different parameters were used to initialise COBYLA: sequential parameters for $p_{\text{target}} = 7$ MaxCut (top row), $\text{LR}_{+\beta}$ parameters for $p_{\text{target}} = 21$ MaxCut (second row) and Max3Sat (bottom row), and $\text{LR}_{-\beta}$ parameters for VertexCover (third row).

also achieve a better residual energy at lower depth, but may be outperformed by some of the "slower" converging parameters at higher depth. This indicates that faster convergence does not necessarily mean better performance. Our proposed sequential method converges most rapidly to a γ-invariant state. Once reached, however, the residual energy remains fixed across subsequent depths, rendering additional QAOA layers ineffective. By contrast, COBYLA initialised with sequential parameters attains a γ-invariant landscape only at p_{target}, thereby enabling lower residual energies. For linear-ramp parameters, the optimisation landscape does not converge to a γ-invariant form at low depth ($p_{\text{target}} = 7$ in our experiments), but does so at higher depth ($p_{\text{target}} = 21$), thereby achieving lower residual energies than the sequential method. This behaviour indicates that slower convergence at greater depth can yield superior results. Hence, the optimal convergence rate depends on the available circuit depth and should be chosen to ensure that convergence coincides with the final QAOA layer.

The rate of change in the optimisation landscape for a given set of parameters depends not only on the absolute values of γ and β, but also on their ratio. We observe that a gradually increasing γ induces a more pronounced vertical shift of the minima towards the region $\beta = 0$ than when γ is initially large and continues to increase. With slowly increasing γ, the extrema broaden more gradually in the horizontal direction of the parameter landscape when β decreases slowly towards zero. In other words: (i) the rate of propagation under the phase operator governs the speed at which the minima converge towards the centre of the landscape, and (ii) the rate of decrease in β determines the speed at which the landscape becomes horizontally uniform (*i.e.*, γ-invariant).

Across methods, the sequential approach performs comparably to, and in some cases better than, alternative strategies, depending on problem and circuit depth. Its drawback lies in the diminishing rate of improvement in residual energy as p increases: At large depth, it is surpassed by methods such as linear ramps that continue to yield improvements throughout the available depths.

In contrast, linear ramps underperform at shallow depth but continue to improve up to the target depth, with the following exceptions: For some problems, notably Vertex-Cover, linear-ramp parameters may fail to yield improve-

ments as p increases. Moreover, for linear ramps to drive the system towards a low-energy state, the mixer ramp must be aligned with both the initial state and the implementation of H_C. In our setting this requires the initial state $|+\rangle$ and a problem Hamiltonian H_C whose minimal eigenvalue encodes the optimum. This corresponds to a negative $\Delta\beta$ (*i.e.*, the $\mathrm{LR}_{-\beta}$ method), since $|+\rangle$ is the ground state of $-U_M$. In contrast, a positive $\Delta\beta$ (*i.e.*, the $\mathrm{LR}_{+\beta}$ method) corresponds to U_M, for which $|+\rangle$ is the *maximum* eigenstate, thereby propagating the system towards a high-energy state. COBYLA performance is highly sensitive to the choice of initialisation: parameters that steer the system towards low-energy states provide a favourable starting point, whereas unfavourable choices, such as the $\mathrm{LR}_{+\beta}$ ramp, are less effective. At shallow depth and for small problem sizes, the optimiser may still identify reasonably good parameters, typically by adjusting mainly the lowest- and highest-depth components of β and γ. At greater depth, however, poor initialisation can prevent convergence to low energies. Although COBYLA results often exhibit substantial variance, at the target depth the residual energies across instances are generally consistent—except in cases of unfavourable initialisation, where performance can deteriorate.

For the problems considered in this paper, the best identified parameters frequently deviate from the patterns reported in the literature. This observation is consistent with the convergence behaviour of the optimisation landscape: As the residual energy becomes increasingly invariant under γ when β approaches zero, larger deviations from the expected γ-patterns can be tolerated without loss in residual energy. If parameters exist for a given p that follow the patterns and approximate an eigenstate of H_C within that depth, which renders the landscape γ-invariant at higher depths, then many other parameter choices may perform comparably well despite substantial deviations from the patterns in their higher-depth components. Furthermore, since the adiabatic condition (*i.e.*, propagation is sufficiently slow so as not to disturb adiabatic dynamics) is most restrictive near the minimum spectral gap, parameters may also deviate substantially from the expected patterns in the low- or intermediate-depth components. Deviations in regions where the spectral gap is large may allow for faster propagation towards the target state without degrading performance. However, since such parameters still require non-arbitrary values of β and γ to drive the system towards a low-energy state, they are unlikely to deviate as strongly from the patterns as the higher-depth, γ-invariant components. Rather, we expect deviations to arise in the smoothness of the increase or decrease, rather than in the overall continuity of the slope of the parameters.

VII. CONCLUSION

We conducted a systematic empirical study of different parametrisation methods in QAOA and their impact on the corresponding optimisation landscapes. Our results challenge the prevailing assumption that optimal parameters consistently follow smooth trends, with β monotonically decreasing and γ monotonically increasing with circuit depth. We find that such patterns are pronounced only at shallow depth. As p increases and β approaches zero, deviations in the higher-depth components of γ become increasingly inconsequential, until in the limit $\beta = 0$ the choice of γ is effectively arbitrary. Iterative parameter fixing yielded consistent performance across all instances, despite not optimising over *all* components of β, γ instead of the full $2p$. While its effectiveness saturates once β approaches 0, the reduced optimisation complexity and strong low-depth performance render it a practical baseline for comparison with more advanced parameter-selection strategies, particularly as we only need to tune 2 variational parameters in each iteration.

While we could reproduce some observations of previous work concerning the performance of different parameter selection methods, we could identify circumstances under which these methods may fail: (1) Linear ramps generally achieve comparable performance across problem instances at high depths, but can be outperformed by alternative methods at shallow depths. Instance-*independent* ramps may lead to marked reductions in solution quality relative to iteratively fixed or optimised parameters. If ramps are misaligned with the initial state or with the implementation of the mixer and phase unitaries, they approximate the solution of the inverse problem (*e.g.*, maximisation instead of minimisation). (2) Exhaustively optimised QAOA parameters achieve best performance among the methods considered, but exhibit substantially higher computational cost than fixed parameters or linear ramps. Results exhibit considerable variance and are highly sensitive to the choice of initialisation. Poor initial conditions often hinder matching the performance of ramps or iterative fixing. Iteratively fixed parameters provide a reliable initialisation strategy.

Our results suggest several promising avenues for improving QAOA. The characteristics outlined in Sec. VI may inform the design of optimiser constraints or motivate variants that iteratively fix parameters with increasing depth. Another direction is to devise fixed linear-ramp methods that adapt to specific problem instances, for example by tuning $\Delta\beta, \Delta\gamma$ through an optimiser or heuristic rather than assuming complete instance independence.

VIII. ACKNOWLEDGEMENT

We acknowledge support from German Federal Ministry of Research, Technology and Space (BMFTR), funding program “Quantum Technologies—from Basic Research to Market”, grant #13NI6092 (VE, MF and WM). WM acknowledges support by the High-Tech Agenda Bavaria.

References

[1] Rajeev Acharya et al. “Quantum error correction below the surface code threshold”. In: *Nature* 638 (2025), pp. 920–926. DOI: 10.1038/s41586-024-08449-y.

[2] Tameem Albash and Daniel A. Lidar. “Adiabatic quantum computation”. In: *Rev. Mod. Phys.* 90 (1 Jan. 2018), p. 015002. DOI: 10.1103/RevModPhys.90.015002.

[3] Kishor Bharti et al. “Noisy intermediate-scale quantum algorithms”. In: *Rev. Mod. Phys.* 94 (1 Feb. 2022), p. 015004. DOI: 10.1103/RevModPhys.94.015004.

[4] Lennart Bittel and Martin Kliesch. “Training Variational Quantum Algorithms Is NP-Hard”. In: *Physical Review Letters* 127.12 (Sept. 2021). Publisher: American Physical Society, p. 120502. DOI: 10.1103/PhysRevLett.127.120502.

[5] Kostas Blekos et al. “A review on Quantum Approximate Optimization Algorithm and its variants”. In: *Physics Reports*. A review on Quantum Approximate Optimization Algorithm and its variants 1068 (June 2024), pp. 1–66. DOI: 10.1016/j.physrep.2024.03.002.

[6] Fernando G. S. L. Brandao et al. *For Fixed Control Parameters the Quantum Approximate Optimization Algorithm’s Objective Function Value Concentrates for Typical Instances*. arXiv:1812.04170 [quant-ph]. Dec. 2018. DOI: 10.48550/arXiv.1812.04170.

[7] Sergey Bravyi et al. “Obstacles to Variational Quantum Optimization from Symmetry Protection”. In: *Phys. Rev. Lett.* 125 (26 Dec. 2020), p. 260505. DOI: 10.1103/PhysRevLett.125.260505.

[8] Cecilia Carbonelli et al. “Challenges for Quantum Software Engineering: An Industrial Application Scenario Perspective”. In: *Quantum Software: Aspects of Theory and System Design*. Ed. by Iaakov Exman et al. Cham: Springer Nature Switzerland, 2024, pp. 311–335. DOI: 10.1007/978-3-031-64136-7_12.

[9] M. Cerezo et al. “Variational Quantum Algorithms”. In: *Nature Reviews Physics* 3.9 (Aug. 2021), pp. 625–644. DOI: 10.1038/s42254-021-00348-9.

[10] Daniel J. Egger, Jakub Mareček, and Stefan Woerner. “Warm-starting quantum optimization”. In: *Quantum* 5 (June 2021), p. 479. DOI: 10.22331/q-2021-06-17-479.

[11] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. *A Quantum Approximate Optimization Algorithm*. arXiv:1411.4028 [quant-ph]. Nov. 2014.

[12] Edward Farhi and Aram W. Harrow. *Quantum Supremacy through the Quantum Approximate Optimization Algorithm*. arXiv:1602.07674 [quant-ph]. Oct. 2019.

[13] Mario Fernández-Pendás et al. “A study of the performance of classical minimizers in the Quantum Approximate Optimization Algorithm”. In: *Journal of Computational and Applied Mathematics* 404 (Apr. 2022), p. 113388. DOI: 10.1016/j.cam.2021.113388.

[14] Gacon, Julien Sebastian. “Scalable Quantum Algorithms for Noisy Quantum Computers”. en. In: (2024). DOI: 10.5075/EPFL-THESIS-11132.

[15] Alexey Galda et al. *Transferability of optimal QAOA parameters between random graphs*. Los Alamitos, CA, USA, Oct. 2021. DOI: 10.1109/QCE52317.2021.00034.

[16] Bartłomiej Gardas et al. “Defects in Quantum Computers”. en. In: *Scientific Reports* 8.1 (Mar. 2018). Publisher: Nature Publishing Group, p. 4539. DOI: 10.1038/s41598-018-22763-2.

[17] Martin Gogeissl, Hila Safi, and Wolfgang Mauerer. “Quantum Data Encoding Patterns and their Consequences”. In: *Proceedings of the 1st Workshop on Quantum Computing and Quantum-Inspired Technology for Data-Intensive Systems and Applications*. Q-Data ’24. Santiago, AA, Chile: Association for Computing Machinery, 2024, pp. 27–37. DOI: 10.1145/3665225.3665446.

[18] Stuart Hadfield et al. “From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz”. In: *Algorithms* 12.2 (2019). DOI: 10.3390/a12020034.

[19] Zichang He et al. “Alignment between initial state and mixer improves QAOA performance for constrained optimization”. en. In: *npj Quantum Information* 9.1 (Nov. 2023). Publisher: Nature Publishing Group, pp. 1–11. DOI: 10.1038/s41534-023-00787-5.

[20] Nishant Jain et al. “Graph neural network initialisation of quantum approximate optimisation”. In: *Quantum* 6 (Nov. 2022), p. 861. DOI: 10.22331/q-2022-11-17-861.

[21] Ali Javadi-Abhari et al. *Quantum computing with Qiskit*. 2024. DOI: 10.48550/arXiv.2405.08810.

[22] Richard M. Karp. “Reducibility among Combinatorial Problems”. In: *Complexity of Computer Computations*. Springer US, 1972, pp. 85–103. DOI: 10.1007/978-1-4684-2001-2_9.

[23] Youngseok Kim et al. “Evidence for the utility of quantum computing before fault tolerance”. en. In: *Nature* 618.7965 (June 2023). Publisher: Nature Publishing Group, pp. 500–505. DOI: 10.1038/s41586-023-06096-3.

[24] Vladimir Kremenetski et al. *Quantum Alternating Operator Ansatz (QAOA) Phase Diagrams and Applications for Quantum Chemistry*. arXiv:2108.13056 [quant-ph]. Oct. 2021. DOI: 10.48550/arXiv.2108.13056.

[25] Tom Krüger and Wolfgang Mauerer. *Out of the Loop: Structural Approximation of Optimisation Landscapes and non-Iterative Quantum Optimisation*. 2024. DOI: 10.48550/arXiv.2408.06493.

[26] Tom Krüger and Wolfgang Mauerer. “Quantum Annealing-Based Software Components: An Experimental Case Study with SAT Solving”. In: *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*. ICSEW’20. Seoul, Republic of Korea: Association for

Computing Machinery, 2020, pp. 445–450. DOI: 10.1145/3387940.3391472.

[27] Xinwei Lee et al. “A Depth-Progressive Initialization Strategy for Quantum Approximate Optimization Algorithm”. en. In: *Mathematics* 11.9 (Jan. 2023). Number: 9 Publisher: Multidisciplinary Digital Publishing Institute, p. 2176. DOI: 10.3390/math11092176.

[28] Xinwei Lee et al. “Parameters Fixing Strategy for Quantum Approximate Optimization Algorithm”. en. In: *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*. Oct. 2021. DOI: 10.1109/QCE52317.2021.00016.

[29] Phillip C. Lotshaw et al. “Empirical performance bounds for quantum approximate optimization”. In: *Quantum Information Processing* 20.12 (Dec. 2021), p. 403. DOI: 10.1007/s11128-021-03342-3.

[30] Wolfgang Mauerer and Stefanie Scherzinger. “1-2-3 Reproducibility for Quantum Software Experiments”. In: *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*. 2022, pp. 1247–1248. DOI: 10.1109/SANER53432.2022.00148.

[31] David Mitchell, Bart Selman, and Hector Levesque. “Hard and easy distributions of SAT problems”. In: *Proceedings of the tenth national conference on Artificial intelligence*. AAAI’92. San Jose, California: AAAI Press, July 1992, pp. 459–465.

[32] J. A. Montañez-Barrera and Kristel Michielsen. “Toward a linear-ramp QAOA protocol: evidence of a scaling advantage in solving some combinatorial optimization problems”. In: *npj Quantum Information* 11.1 (Aug. 2025). DOI: 10.1038/s41534-025-01082-1.

[33] J. A. Montañez-Barrera, Dennis Willsch, and Kristel Michielsen. “Transfer learning of optimal QAOA parameters in combinatorial optimization”. In: *Quantum Information Processing* 24.5 (May 2025). DOI: 10.1007/s11128-025-04743-4.

[34] Maniraman Periyasamy et al. “Guided-SPSA: Simultaneous Perturbation Stochastic Approximation Assisted by the Parameter Shift Rule”. In: *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*. Vol. 01. 2024, pp. 1504–1515. DOI: 10.1109/QCE60285.2024.00177.

[35] Alberto Peruzzo et al. “A variational eigenvalue solver on a quantum processor”. In: *Nature Communications* 5.1 (July 2014). arXiv:1304.3061 [physics, physics:quant-ph], p. 4213. DOI: 10.1038/ncomms5213.

[36] M. J. D. Powell. “A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation”. en. In: Book Title: Advances in Optimization and. Springer Netherlands, 1994. DOI: 10.1007/978-94-015-8330-5_4.

[37] Hila Safi, Karen Wintersperger, and Wolfgang Mauerer. “Influence of HW-SW-Co-Design on Quantum Computing Scalability”. In: *2023 IEEE International Conference on Quantum Software (QSW)*. 2023, pp. 104–115. DOI: 10.1109/QSW59989.2023.00022.

[38] Lukas Schmidbauer et al. *Path Matters: Industrial Data Meet Quantum Optimization*. 2025. DOI: 10.48550/arXiv.2504.16607.

[39] Manuel Schönberger, Immanuel Trummer, and Wolfgang Mauerer. “Quantum-Inspired Digital Annealing for Join Ordering”. In: *Proc. VLDB Endow.* 17.3 (Nov. 2023), pp. 511–524. DOI: 10.14778/3632093.3632112.

[40] Michael Streif and Martin Leib. “Training the quantum approximate optimization algorithm without access to a quantum processing unit”. In: *Quantum Science and Technology* 5.3 (May 2020), p. 034008. DOI: 10.1088/2058-9565/ab8c2b.

[41] James Sud et al. “Parameter-setting heuristic for the quantum alternating operator ansatz”. In: *Phys. Rev. Res.* 6 (2 May 2024), p. 023171. DOI: 10.1103/PhysRevResearch.6.023171.

[42] Reuben Tate et al. “Warm-Started QAOA with Custom Mixers Provably Converges and Computationally Beats Goemans-Williamson’s Max-Cut at Low Circuit Depths”. In: *Quantum* 7 (Sept. 2023), p. 1121. DOI: 10.22331/q-2023-09-26-1121.

[43] Simone Tibaldi et al. “Bayesian Optimization for QAOA”. In: *IEEE Transactions on Quantum Engineering* 4 (2023). Conference Name: IEEE Transactions on Quantum Engineering. DOI: 10.1109/TQE.2023.3325167.

[44] V Vijendran et al. “An expressive ansatz for low-depth quantum approximate optimisation”. In: *Quantum Science and Technology* 9.2 (Feb. 2024), p. 025010. DOI: 10.1088/2058-9565/ad200a.

[45] Jonathan Wurtz and Peter Love. *Classically optimal variational quantum algorithms*. 2021. DOI: 10.1109/TQE.2021.3122568.

[46] Jonathan Wurtz and Peter Love. “MaxCut quantum approximate optimization algorithm performance guarantees for $p>1$”. In: *Physical Review A* 103.4 (Apr. 2021). Publisher: American Physical Society, p. 042612. DOI: 10.1103/PhysRevA.103.042612.

[47] Jonathan Wurtz and Peter J. Love. “Counterdiabaticity and the quantum approximate optimization algorithm”. In: *Quantum* 6 (Jan. 2022), p. 635. DOI: 10.22331/q-2022-01-27-635.

[48] Jonathan Wurtz and Danylo Lykov. *Fixed-angle conjectures for the quantum approximate optimization algorithm on regular MaxCut graphs*. Nov. 2021. DOI: 10.1103/PhysRevA.104.052419.

[49] Leo Zhou et al. “Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices”. In: *Physical Review X* 10.2 (June 2020). Publisher: American Physical

Society, p. 021067. DOI: 10.1103/PhysRevX.10.021067.

APPENDIX A. LANDSCAPE EVALUATION ALGORITHM

Algorithm 1 Landscape scan with iteratively increasing depth

Input: $\mathcal{P}$ = {problem type, problem size, graph degree, clause to variable ratio, seed} - parameters for problem specification;
$\mathcal{S}$ = {resolution,β bounds, γ bounds} - parameters for scan grid specification;
$\mathcal{I}$ = $\{p_{\text{start}}, p_{\text{target}}, p_{\text{step}}\}$ - parameters for iteration loop specification;
$\vec{\gamma_{\text{init}}}, \vec{\beta_{\text{init}}}$ - optional set of parameters which will be fixed instead of the best found values;
opt - optimal optimiser with which to optimise $\vec{\gamma_{\text{init}}}, \vec{\beta_{\text{init}}}$ before starting the scan;
Output: $\mathcal{E}$ - List of grids of expectation values corresponding to each G_p evaluated at each iteration

```
H_C ← mapProblemToIsing(P)
if γ_init, β_init, opt ≠ None then
    res ← QAOA(H_C, p_target, γ_init, β_init, opt)
            .computeMinimumEigenvalue()
    γ_init, β_init ← res.opt_parameters
end if
γ_b, β_b ← None
E ← List()
p ← p_start
while p ≤ p_target do
    ansatz ← QAOAAnsatz(H_C, p)
    G_p ← makeGridofParams(S, p, γ_b, β_b)
    G_e ← Estimator(ansatz, G_p, shots = None)
    if γ_init, β_init ≠ None then
        γ_b, β_b ← γ_init[0 : p], β_init[0 : p]
    else
        x, y ← getIndex(min(G_e))
        γ_b, β_b ← G_p[x][y]
    end if
    E.append(G_e)
    p ← p + p_step
end while
```

Algorithm 1 lists the procedure for scanning the optimisation landscapes using different parameter selection methods. The scan starts by initialising the problem and, in the case that an optimiser has been set, optimises parameters for the target p_{target} starting from given initial parameters. In the case that no optimiser has been set, this step is skipped. Then, starting from the given start depth, lower depth parameters are fixed to those of the previous best parameters, the initial parameters for LR, or the optimised parameters if an optimiser is used. A grid of parameters within the given bounds is then created using Algorithm 2. Subsequently, a QAOA ansatz from the cost operator and current depth is created. We then pass the ansatz, cost operator and the grid of parameters to the Qiskit `Estimator`. The `Estimator` then binds each of the parameter sets in the grid to a copy of the ansatz and calculates the expectation value of the cost operator in the state prepared by this circuit. As a result we get a grid of expectation values of the same dimensions as out grid of parameter sets, which we can plot as a landscape. The best expectation values for each p are used to compute the approximation value plots.

Algorithm 2 makeGridofParams

Input: n - resolution of grid;
$B_{\{\text{upper, lower}\}}(\{\beta,\gamma\})$ - {upper,lower} bound on $\{\beta,\gamma\}$;
p - circuit depth/length of parameter-vectors in the grid;
$\vec{\gamma_f}$ - the γ parameters to be fixed at lower depths;
$\vec{\beta_f}$ - the β parameters to be fixed at lower depths;
Output: G_p - Grid of size $n \times n$ with the points in the grid corresponding to a tuple of parameter-vectors $(\vec{\gamma}, \vec{\beta})$ where each parameter-vector has length p

```
G_p ← List(shape = (n, n)
intervals_β ← B_upper(β) − (i/n)(B_upper(β) − B_lower(β)) for i in range(n)
intervals_γ ← B_upper(γ) − (i/n)(B_upper(γ) − B_lower(γ)) for i in range(n)
i ← 0
while i ≤ n do
    j ← 0
    while j ≤ 0 do
        β_ij ← β_f + intervals_β[i]
        γ_ij ← γ_f + intervals_γ[j]
        G_p[i][j] ← (β_ij, γ_ij)
        j ← j + 1
    end while
    i ← i + 1
end while
```

APPENDIX B. MAX3SAT – "EASY" INSTANCES

In order to determine how the landscape changes when when we choose a different clause to variable ratio, which is not in the range generally considered as hard ($\alpha \in (3.5; 4.9]$), we determined the landscapes resulting from QAOA applied to 10 Max3SAT instances with $\alpha \notin (3.5; 4.9]$, which we will refer to as "easy" instances. The $\bar{r}$ landscape is shown in Fig. 16 and the σ_r landscape is shown in Fig. 17. In both figures, the facets corresponding to the different methods are arranged in the same way as for the hard instance plots (Fig. 12 and Fig. 13). The average approximation quality of the energies produced by the parameters given by these methods is shown in Fig. 18.

The $\bar{r}$ landscapes produced by the different methods (see Fig. 16) progress similarly to those of the hard instances, but with a generally lower σ_r (see Fig. 17), especially for the landscapes given by the parameters optimised by COBYLA (initialised with $\text{LR}_{+\beta}$).

Notably, all methods perform similarly as in the case where we chose instances from the (hard) range $\alpha \in$

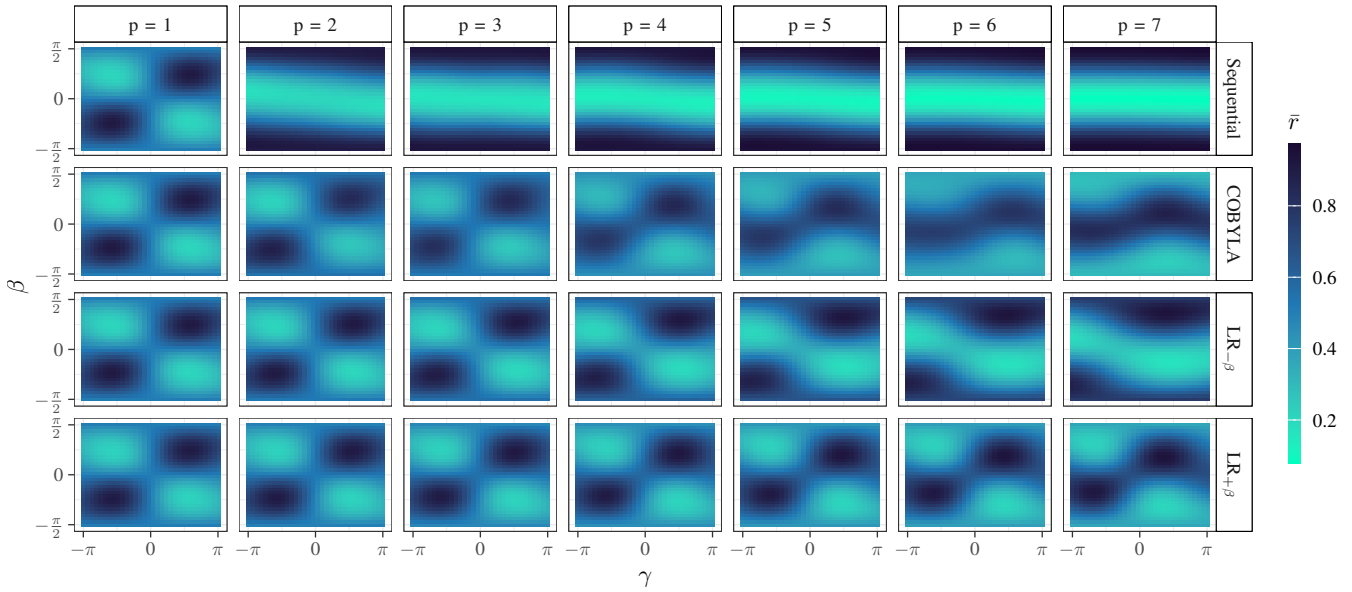

FIGURE 16: Average residual energy $\bar{r}$ for 10 Max3SAT instances of sizes 14 to 24, with $\alpha \notin (3.5; 4.9]$ and γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

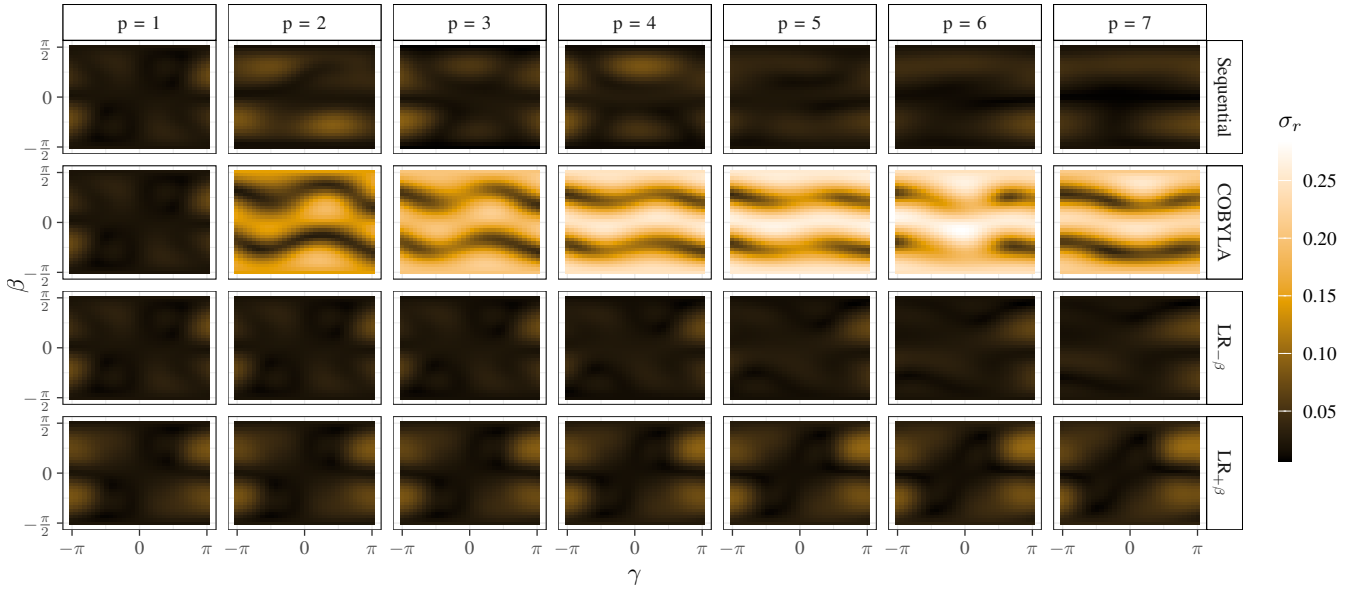

FIGURE 17: Standard deviation of the residual energy $\bar{r}$ in Fig. 16 over 10 Max3SAT instances.

$(3.5; 4.9]$, the main difference is the magnitude of σ_r.

APPENDIX C. TYPICAL SINGLE INSTANCE SCANS AND PARAMETER SHAPE

In Sec. V we focus on the average and standard deviation of the residual energy over multiple instances. This appendix supplements the corresponding plots for typical single instances of problems examined in that section, to provide a frame of reference for those results. For the energy landscape scans, we plot the energy expectation value $F_p(\gamma, \beta)$.

A. MAXCUT

Fig. 19 and Fig. 20 show the results for one of the MaxCut instances examined in Sec. V-B. The results for the same MaxCut instance, but at a higher depth (*i.e.*, $p_{\text{target}} = 21$) and incrementing p in steps of 2, are shown in Fig. 21 and Fig. 22.

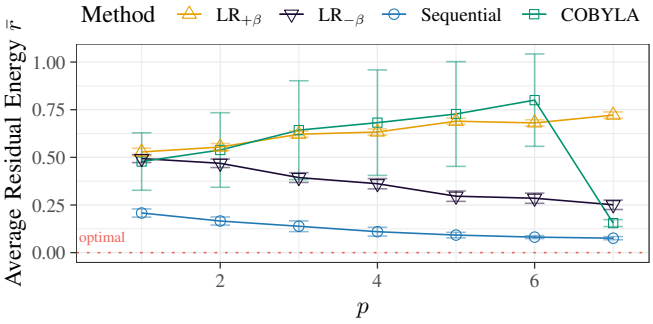

FIGURE 18: Average and standard deviation of the approximation quality for 10 Max3SAT instances of sizes 14 to 24 with $\alpha \notin (3.5; 4.9]$ when fixing parameters at each layer p according to the considered methods.

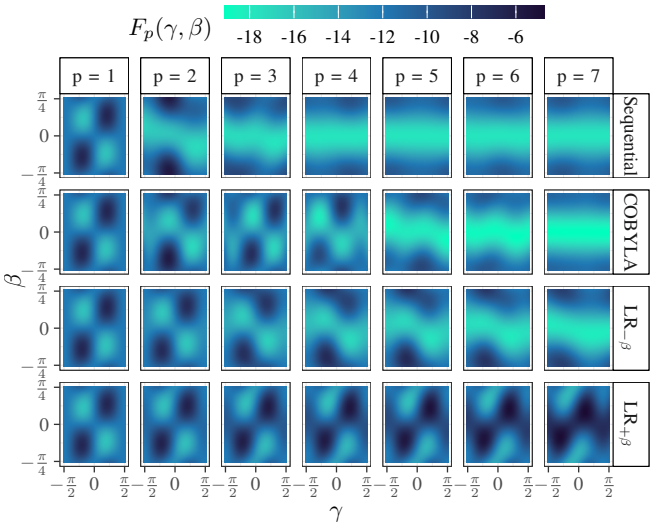

FIGURE 19: Energy expectation landscapes for one MaxCut instance on a 3-regular 16-vertex graph with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

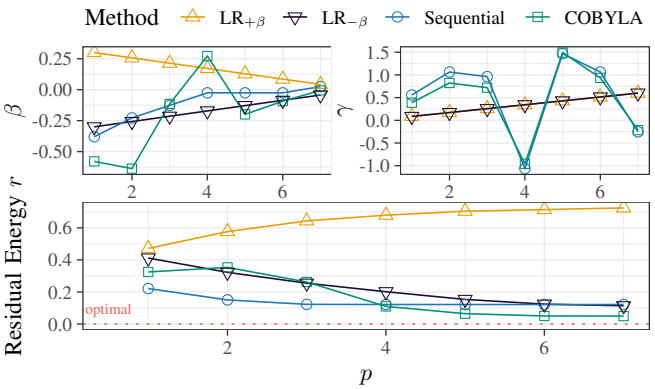

FIGURE 20: Approximation quality and arrangement of the parameters fixed at each QAOA layer for the MaxCut instance and methods shown in Fig. 19.

B. VERTEXCOVER

Fig. 23 and Fig. 24 correspond to a single instance of the VertexCover problem examined in Sec. V-D. Notably, the density of local minima in the $p = 1$ landscape is even more pronounced than in the average landscape shown in Fig. 9.

C. MAX3SAT

Fig. 25 and Fig. 26 show the results for a typical Max3SAT instance from the "hard" range examined in Sec. V-E.

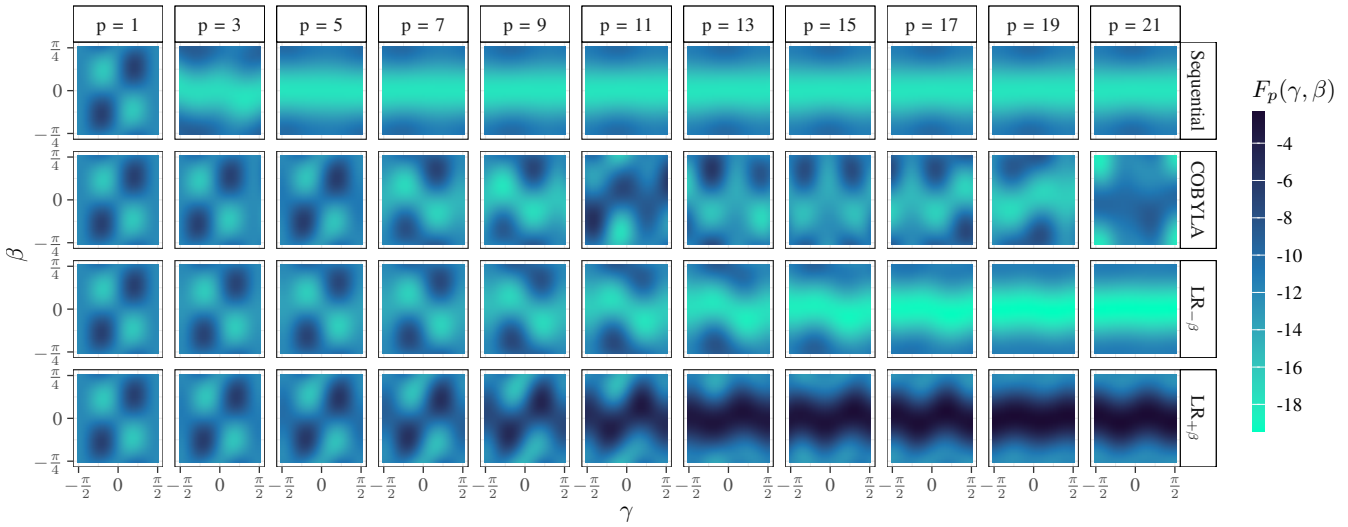

FIGURE 21: Energy landscapes for the same 3-regular 16-vertex graph instance shown in Fig. 19 with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

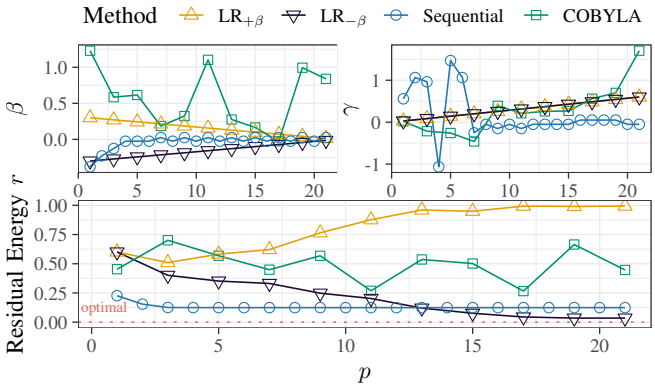

FIGURE 22: Approximation quality and arrangement of the parameters fixed at each QAOA layer for the MaxCut instance and methods shown in Fig. 21.

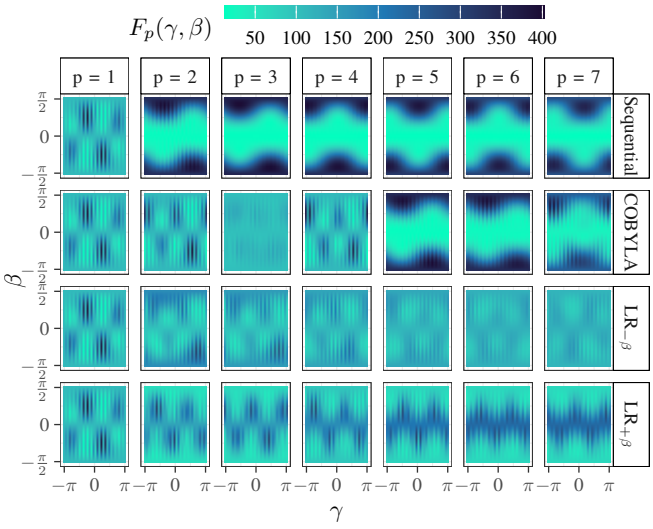

FIGURE 23: Energy landscapes for one VertexCover instance on a 3-regular 16-vertex graph with γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

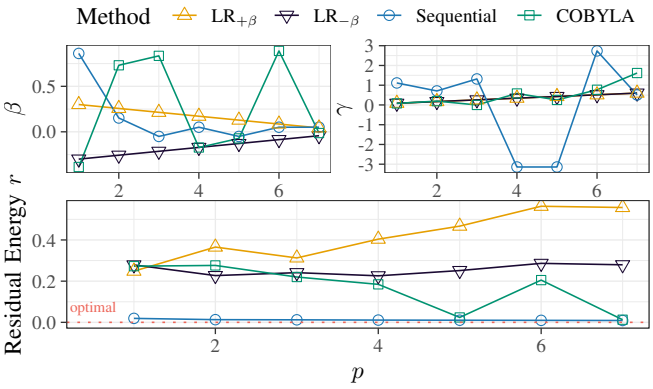

FIGURE 24: Approximation quality and arrangement of the parameters fixed at each QAOA layer for the VertexCover instance and methods shown in Fig. 23.

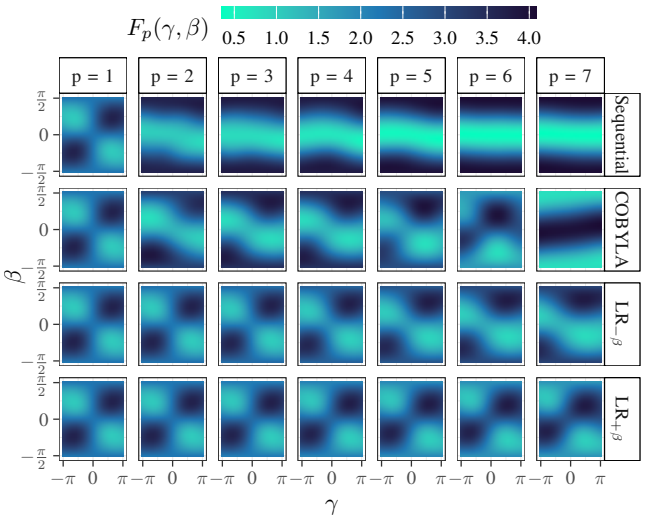

FIGURE 25: Energy landscapes for one 21-qubit Max3SAT instance with $\alpha = 4.25 \in (3.5; 4.9]$ and γ, β set to sequentially fixed parameters (top row), optimised parameters using COBYLA starting from the sequential parameters (second row), LR$_{-\beta}$ parameters with $\Delta_\beta = -0.3, \Delta_\gamma = 0.6$ (third row), and LR$_{+\beta}$ parameters with $\Delta_\beta = 0.3, \Delta_\gamma = 0.6$ (bottom row).

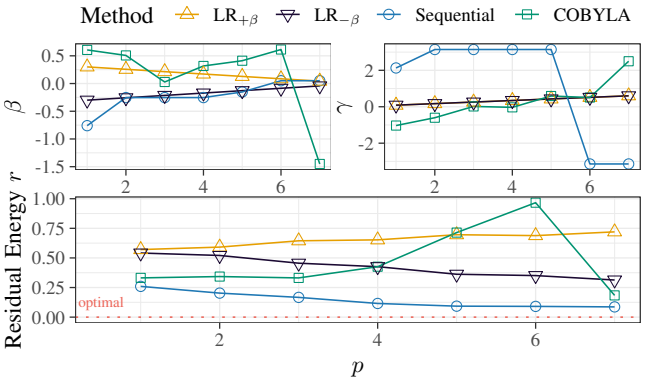

FIGURE 26: Approximation quality and arrangement of the parameters fixed at each QAOA layer for the Max3SAT instance and methods shown in Fig. 25.