There is no unique way to encode a quantum algorithm into a quantum circuit. With limited qubit counts, connectivity, and coherence times, a quantum circuit optimization is essential to make the best use of near-term quantum devices. We introduce a new circuit optimizer called Aqcel, which aims to remove redundant controlled operations from controlled gates, depending on initial states of the circuit. Especially, the Aqcel can remove unnecessary qubit controls from multi-controlled gates in polynomial computational resources, even when all the relevant qubits are entangled, by identifying zero-amplitude computational basis states using a quantum computer. As a benchmark, the Aqcel is deployed on a quantum algorithm designed to model final state radiation in high energy physics. For this benchmark, we have demonstrated that the Aqcel-optimized circuit can produce equivalent final states with much smaller number of gates. Moreover, when deploying Aqcel with a noisy intermediate scale quantum computer, it efficiently produces a quantum circuit that approximates the original circuit with high fidelity by truncating low-amplitude computational basis states below certain thresholds. Our technique is useful for a wide variety of quantum algorithms, opening up new possibilities to further simplify quantum circuits to be more effective for real devices.

1 Introduction

Recent technology advances have resulted in a variety of universal quantum computers that are being used to implement quantum algorithms. However, these noisy-intermediate-scale quantum (NISQ) devices [1] may not have sufficient qubit counts, qubit connectivity and capability to stay coherent for the entirety of operations in a particular algorithm implementation. Despite these challenges, a variety of applications have emerged across science and industry. For example, there are many promising studies in experimental and theoretical high energy physics (HEP) for exploiting quantum computers. These studies include event classification [2, 3, 4, 5, 6, 7], reconstructions of charged particle trajectories [8, 9, 10, 11] and physics objects [12, 13], unfolding measured distributions [14] as well as simulation of multi-particle emission processes [15, 16]. A common feature of all of these algorithms is that only simplified versions can be run on existing hardware due to the limitations mentioned above.

There are generically two strategies for improving the performance of NISQ computers to execute existing quantum algorithms. One strategy is to mitigate errors through active or passive modifications to the quantum state preparation and measurement protocols. For example, readout errors can be mitigated through post-processing steps [17, 18, 19, 20, 21, 22] and gate errors can be mitigated by systematically enlarging errors before extrapolating to zero error [23, 24, 25, 26, 27, 28]. A complementary strategy to error mitigation is quantum compilation. There is no unique way to encode a quantum algorithm into a set of gates, and certain
realizations of an algorithm may be better suited for a given quantum device. Widely used tools are Qiskit [29] and t(ket) [30], which contain a variety of architecture-agnostic and architecture-specific routines. There are also a variety of other toolkits for circuit optimization, including hardware-specific packages for quantum circuits [31, 32, 33, 34, 35, 29, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49].

Among the gates used for an algorithm encoding, multi-controlled gates are significant error sources because they result in many CNOT gates after the decomposition [50], and also require SWAP gates to fit within limited qubit topology. The costs to implement multi-controlled gates can be reduced by using relative phase Toffoli gates [51], ancilla qubits [52] or qutrits [53, 54, 55, 56, 57, 58, 59] in the implementation.

An alternative approach for reducing the costs is to remove unnecessary qubit controls. The reversible circuit synthesis can reduce redundant qubit controls while maintaining the equivalence of a quantum circuit before and after the optimization [60]. Previous work with reversible circuit synthesis has largely focused on circuits composed of CNOT gates [61]. The circuit synthesis is extended later to more general quantum circuits, e.g., those composed of CNOTs and Z-basis rotation gates [62]. It is possible to remove, beyond reversible circuit synthesis, more unnecessary controlled operations if we consider maintaining the equivalence of the final state. Imagine that there is a $n$-qubit quantum circuit designed to work with different initial states and it is executed with a given initial state such as $|0\rangle^{\otimes n}$. In this case, the circuit will reach only a selected set of intermediate states and some operations may become trivial. Such initial-state dependent circuit optimization may find more rooms for optimization if the equivalence of the final state, not the circuit itself, is preserved. Thus, it will enable more aggressive reduction of unnecessary controlled operations than initial-state independent circuit optimization.

There are two main approaches for initial-state dependent circuit optimization. The first one is a continuous optimization that trains an ansatz with parameters [63, 64]. The second one is a discretized optimization in which some controlled operations are removed from a gate if the quantum state satisfies a specific condition at the point where the gate is operated [49, 65, 66]. We focus on the latter in this paper. Among existing discretized optimization protocols that account for initial states, the Relaxed Peephole Optimization (RPO) [49] reduces controlled operations when the qubits in the X- or Z-basis states are used as control qubits or a target qubit of a controlled gate. This protocol, however, cannot remove qubit controls in the case where all relevant qubits are entangled. The ZX-calculus [65, 66] exploits a copy-rule for removing a qubit control from CNOT gate when the control qubit state is $|0\rangle$. This leads to an initial-state dependent circuit optimization, but it cannot remove all qubit controls within polynomial complexity.

The novel optimization protocol proposed in this paper has three distinct features. First, there is the ability of removing redundant qubit controls no matter whether all the relevant qubits are entangled or not. Second, the identification of zero- or low-amplitude computational basis states using a quantum computer allows one to obtain bitstrings in polynomial time, otherwise exponential resources are required in the classical calculations. Third, the decomposition of multi-controlled $U$ gates into Toffoli gates and singly-controlled $U$ gates enables us to perform the search of all unnecessary qubit controls in polynomial resources. This new optimization protocol also serves a new efficient method to approximate quantum circuits in the NISQ era by truncating low-amplitude computational basis states that do not contribute significantly to the final state.

This optimization protocol is called AQCEL (and pronounced “excel”) for Advancing Quantum Circuit by iCEPP and LBNL. To demonstrate the effectiveness of the AQCEL protocol, we will use a quantum algorithm that models a parton shower [16]. This algorithm provides a useful benchmark because it is designed to work with different initial states corresponding to different initial particles, meaning that the quantum circuits have redundancy for a specific initial state.

This paper is organized as follows. Section 2 provides an overview of the AQCEL protocol. The

---

1 In fact, the ZX-calculus is complete in the formal logic sense of the word, such that one can always prove that all unnecessary qubit controls can be removed using rules of the ZX-calculus [67]. However, in general this scheme requires exponential resources. Nevertheless, the ZX-calculus is still incredibly powerful and underlies many of the optimization techniques of quantum transpilers.
application of this protocol to the HEP example is presented in Sec. 3. Following a brief discussion about the applicability and future extensions of the protocol in Sec. 4, the paper concludes in Sec. 5.

2AQCEL optimization protocol

First, we summarize the concept of removing redundant controlled operations from controlled gates depending on initial states of a quantum circuit. Then, the AQCEL protocol of removing redundant qubit controls and the methods for executing the whole optimization in polynomial resources are described.

2.1 Basic idea of redundant controlled operations removal

A controlled gate performs a different operation depending on the quantum state at the point where the gate is applied. Let $m$ be the number of control qubits of this gate. Consider expanding the state of the full system $|\psi\rangle$ into a superposition of computational basis states as

$$|\psi\rangle = \sum_{j,k} c_{j,k} |j\rangle_{\text{ctl}} \otimes |k\rangle,$$

where $|j\rangle_{\text{ctl}}$ denotes the state of the control qubits, while the unlabeled ket corresponds to the rest of the system. We write the states as integers with $0 \leq j \leq 2^m - 1$ and $0 \leq k \leq 2^{n-m} - 1$. We assume that the controlled gate is applied to computational basis states whose bitstrings on all control qubits are 1, which corresponds to the state $|j\rangle_{\text{ctl}} = |11 \cdots 1\rangle = |2^m - 1\rangle_{\text{ctl}}$. This allows one to classify the state of the system into three general classes using the amplitudes $c_{j,k}$:

**Triggering**: $c_{j,k} \neq 0$ if and only if $j = 2^m - 1$. The controlled operation of the gate in question is applied for all computational bases in the superposition.

**Non-triggering**: $c_{2^m-1,k} = 0$ for all $k$. The controlled operation is never applied.

**Undetermined**: The state is neither triggering nor non-triggering.

A circuit containing triggering or non-triggering controlled gates can be simplified by removing all controls (triggering case) or by eliminating the gates entirely (non-triggering case).

While an undetermined single-qubit controlled gate cannot be simplified under the current scheme, an undetermined multi-qubit controlled gate can be by removing the controls on some of the qubits, if the state of the system satisfies the condition, and that is our interest.

$$|0\rangle \xrightarrow{H}$$

$$|0\rangle$$

$$|0\rangle$$

Figure 1: A quantum circuit in $|0\rangle^\otimes 3$ initial state. The control and target qubits for the Toffoli gate are entangled because the Hadamard and CNOT gates create the GHZ state.

As an example of this concept, consider the simple circuit in Fig. 1 composed of three qubits in $|0\rangle^\otimes 3$ initial state. At the Toffoli gate, the quantum state is in the superposition of $|000\rangle$ and $|111\rangle$, which is a Greenberger-Horne-Zeilinger state where the qubits are maximally entangled. This is an undetermined state for the Toffoli gate. Moreover, all the control qubits and the target qubit are entangled, which is a difficult case for the qubit control reduction. However, the AQCEL can remove one of the two qubit controls from the Toffoli, replacing it with a CNOT gate controlled only by the remaining one.

2.2 General conditions to eliminate qubit controls

Given a multi-qubit controlled-$U$ gate with $m$ control qubits, denoted by $C^m[U]$, and a system in an undetermined state $|\psi\rangle$ defined in Sec. 2.1, we can derive general conditions for a part of the controlled operations to be removed, as follows.

Let $x$ ($< m$) be the number of controls to be removed. Without loss of generality, the decomposition of $|\psi\rangle$ can be rewritten as

$$|\psi\rangle = \sum_{i,l,k} \tilde{c}_{i,l,k} |i\rangle_{\text{ctl}}' \otimes |l\rangle_{\text{free}} \otimes |k\rangle,$$

where $|i\rangle_{\text{ctl}}'$ and $|l\rangle_{\text{free}}$ are the states of the $m - x$ remaining control qubits and the $x$ qubits from which the controls are removed. From Eq. (1),

$$|i\rangle_{\text{ctl}}' \otimes |l\rangle_{\text{free}} = |2^x i + l\rangle_{\text{ctl}},$$

and therefore

$$\tilde{c}_{i,l,k} = c_{2^x i + l,k}.$$
Applying the original controlled gate to $|\psi\rangle$ yields
\[
C^m[U] |\psi\rangle = \sum_{j=0}^{2^m-2} \sum_k c_{j,k} |j\rangle |k\rangle + \sum_k c_{2^m-1,k} (2^m - 1) U |k\rangle,
\]
where ket subscripts and the tensor product symbols are omitted for simplicity. In contrast, a new gate with fewer controls gives
\[
C^{m-x}[U] |\psi\rangle = \sum_{i=0}^{2^{m-x}-2} \sum_{l,k} \tilde{c}_{i,l,k} |i\rangle |l\rangle |k\rangle + \sum_{l,k} \tilde{c}_{2^{m-x}-1,l,k} (2^{m-x} - 1) |l\rangle U |k\rangle.
\]
For the removal of $x$ qubit controls to be allowed, the right hand sides of Eqs. (5) and (6) must be identical. This requires
\[
\sum_{l=0}^{2^x-2} \sum_k \tilde{c}_{2^{m-x}-1,l,k} (2^{m-x} - 1) |l\rangle U |k\rangle = \sum_{l=0}^{2^{m-x}-2} \sum_k c_{2^{m-x}-2x+l,k} (2^m - 2^x + l) |k\rangle.
\]
Denoting $U |k\rangle = \sum_{k'} u_{k,k'} |k'\rangle$ (replacing $k' \leftrightarrow k$ on the left hand side)
\[
\sum_{l=0}^{2^x-2} \sum_{k,k'} \tilde{c}_{2^{m-x}-1,l,k'} u_{k,k'} (2^{m-x} - 1) |l\rangle |k\rangle = \sum_{l=0}^{2^{m-x}-2} \sum_k c_{2^{m-x}-1,l,k} (2^{m-x} - 1) |l\rangle |k\rangle.
\]
Then, we have
\[
\sum_{k,k'} \tilde{c}_{2^{m-x}-1,l,k'} u_{k,k'} = \tilde{c}_{2^{m-x}-1,l,k}\forall l \in \{0, 1, ..., 2^x - 2\}, k.
\]
Eq. (10) holds if the row vector $\{\tilde{c}_{2^{m-x}-1,l,k}\}$ is an eigenvector of the matrix $u$ with eigenvalue 1 under right multiplication for $0 \leq l \leq 2^x - 2$, or if $\tilde{c}_{2^{m-x}-1,l,k} = 0$ for $0 \leq l \leq 2^x - 2$ and all $k$.

Since the cost of exactly computing the complex amplitudes of the quantum state is exponential, in AQCEL we only consider this second condition:
\[
\tilde{c}_{2^{m-x}-1,l,k} = 0 \forall l \in \{0, 1, ..., 2^x - 2\}, k.
\]
This removal of redundant qubit controls therefore requires us to find out if $|\psi\rangle$ satisfies Eq. (11) for each controlled gate.

In the previous optimization based on RPO, if the target qubit is assumed to be entangled, the condition for the removal of qubit controls from $C^m[U]$ is that control qubits which will be removed must be $|1\rangle$ in a pure state. The condition can be written as
\[
\tilde{c}_{i,l,k} = 0 \forall l \in \{0, 1, ..., 2^x - 2\}, i, k.
\]

Eq. (12) is a sufficient condition of Eq. (11), which means that the AQCEL has more rooms for removal of unnecessary qubit controls. A difference between Eq. (11) and Eq. (12) resides in whether one can remove unnecessary qubit controls from entangled control qubits or not, under the condition that target qubit is also entangled.

### 2.3 Identification of computational basis states

In general, a circuit consisting of $n$ qubits creates a quantum state described by a superposition of all of the $2^n$ computational basis states. However, it is rather common that a specific circuit produces a quantum state where only a subset of the computational basis states has nonzero amplitudes. Moreover, the number of nonzero-amplitude basis states depends on the initial state. This is why the three classes of the states on control qubits arise.

As mentioned in Sec. 2.2, we have to find out if $|\psi\rangle$ satisfies Eq. (11). It can be written down as
\[
\sum_k |\tilde{c}_{2^{m-x}-1,l,k}|^2 = 0 \forall l \in \{0, 1, ..., 2^x - 2\}.
\]
Eq. (13) requires that there is no computational basis state whose bitstring on all control qubits of $C^{m-x}[U]$ is $|11 \cdots 1\rangle$, except when the bitstring on the removed $x$ control qubits is also $|11 \cdots 1\rangle$. In other words, there should be no bitstring by which $C^m[U]$ is not triggered but $C^{m-x}[U]$ is. This can be verified by the identification of bitstrings on all control qubits of $C^m[U]$.

The possible bitstrings on control qubits at each controlled gate can be determined either through a classical simulation or by measuring the control qubits by a quantum computer repeatedly. In the case of a classical simulation, one performs the full calculation of the amplitudes. When instead the quantum measurements...
are used, the circuit is truncated right before the controlled gate in question, and the control qubits are measured repeatedly at the truncation point. Finiteness of the relevant amplitudes can be inferred from the distribution of the obtained bitstrings, albeit within the statistical uncertainty of the measurements.

A few notes should be taken on the computational costs of the two methods. Consider an \( n \)-qubit circuit with \( N \) controlled gates. A classical simulation of the state vector before a given controlled gate has an exponential scaling in the number of qubits and requires \( O(2^n) \) computations. On the other hand, measuring \( m \) control qubits \( M \) times on each controlled gate by a quantum computer only requires \( O(MN^2 + mMN) \) operations which scales only polynomially with the number of qubits. More details on the estimates of the computational resource necessary for the identification of computational basis states are described in Appendix B.

Note that for noisy quantum computers the measurements of the bitstrings will not be exact due to hardware noise. The list of observed bitstrings would contain contributions from errors on the preceding gates and the measurement itself. In AQCEL, we obtain the calibration matrix for the control qubits (with 8192 shots per measurement) using Qiskit Ignis API [29]. The matrix is then applied to the observed distribution with a least-squares fitting approach. To deal with remaining error contributions after the measurement error mitigation, we opt to ignore the observed bitstrings with occurrence below certain thresholds. Once such a threshold has been decided, the number of measurements required has to be large enough for the statistical uncertainty to be smaller than this threshold.

In order to choose the thresholds, we consider gate errors in the single-qubit gates and CNOT gates. Let the single-qubit gate and CNOT error rates be \( \epsilon_u(i) \) and \( \epsilon_{\text{cx}}(i,j) \), respectively, with \( i \) and \( j \) indicating qubits that the gates act on. We can approximate the probabilities, \( p_u \) and \( p_{\text{cx}} \), of measuring the bitstrings without any single-qubit gate or CNOT gate errors occurring anywhere in the circuit by performing qubit-wise (index-dependent) multiplications of the error rates:

\[
p_u = \prod_i \left(1 - \epsilon_u(i)\right)^{n_u(i)},
\]

\[
p_{\text{cx}} = \prod_{i \neq j} \left(1 - \epsilon_{\text{cx}}(i,j)\right)^{n_{\text{cx}}(i,j)},
\]

where \( n_u(i) \) and \( n_{\text{cx}}(i,j) \) are the numbers of single-qubit gates and CNOT gates acting on the corresponding qubits, respectively. The probability \( p_e \) of measuring the bitstrings with at least one gate error occurring anywhere in the circuit is

\[
p_e = 1 - p_u p_{\text{cx}} \sim N_{\text{cx}} \epsilon_{\text{cx}}.
\]

In the last approximation, we have assumed that all CNOT errors are equal, and much larger than single-qubit gate errors but still much smaller than one: \( \epsilon_u(i) \ll \epsilon_{\text{cx}}(i,j) = \epsilon_{\text{cx}} \ll 1 \). From the \( p_e \), the first threshold is chosen to be

\[
s^\text{dyn}_e := p_e / 2^m,
\]

where \( m \) is the number of the measured control qubits. This choice of dynamical threshold is motivated by assuming that the quantum errors would result in a uniform distribution of all possible bitstrings according to a depolarizing error model. It should be noted that the \( p_e \) increases as the circuit execution proceeds because the \( p_e \) accounts for the error rates from all the preceding gates in the circuit. As an alternative strategy to the dynamical threshold, we also examine the static thresholds, \( s^\text{f}_e \), that are kept constant throughout the circuit, with the values between 0.005 and 0.3.

Discarding all bitstrings with occurrence under certain thresholds usually modifies the final state of the optimized circuit from one of the original circuit. On the other hand, applying certain

---

\(^3\)When the number of measurements is not enough for a part of low-amplitude computational basis states, the AQCEL approximates the original circuit to a simpler circuit, but this usually loses the identity of the final state. There is a similar effect when low-amplitude computational basis states below thresholds are truncated, as described later.

\(^4\)In the actual implementation, the threshold of 0.005 is always applied to suppress contributions from imperfect measurement error mitigation.

\(^5\)This is justified under the assumption that the residual gate errors act as a perturbation, inserting spurious computational basis states with small amplitudes into the superposition of the system.

\(^6\)The reported error rates at the time of the experiment, measured during the preceding calibration run of the hardware, are used for the threshold calculation.
thresholds will leave high-amplitude computational basis states, while rejecting low-amplitude computational basis states which do not contribute to the final result meaningfully. Thus, with the proper thresholds, one can produce a quantum circuit which well approximates the original circuit by removing unimportant qubit controls that trigger on low-amplitude computational basis states. In other words, the actual threshold of Aqcel should be selected by considering the trade-off between the noise resilience and the identity of the final state to the original ideal state.

2.4 Elimination of redundant controlled operations

Once the nonzero-amplitude computational basis states are identified at each controlled gate, the next step is to figure out which qubit controls can be removed using Eq. (11). The computational cost of determining the removal of redundant qubit controls would be at most $O(Mn^4a^mN)$ (Appendix B), which scales exponentially with $m$, the number of control qubits ($N$ is the number of multi-qubit controlled gates in the circuit). To avoid this, a generic multi-qubit controlled gate should be decomposed into controlled gates with small, fixed number of control qubits. An arbitrary multi-qubit controlled-$U$ gate with $m$ control qubits can be decomposed into $O(m)$ Toffoli and controlled-$U$ gates [50]. Besides, these Toffoli gates can be replaced with relative phase implementation of Toffoli gates (referred to as just “Toffoli” hereafter) [51], which reduces the CNOT counts from 6 (in the regular Toffoli decomposition) to 3. Therefore, in the Aqcel scheme, we assume that all controlled gates in a quantum circuit are reduced to Toffoli gates denoted as $C^2[X]$ and singly-controlled unitary operations denoted as $C[U]$. This results in a significant reduction of computational cost of the decision of all redundant qubit controls, because all controlled gates have either 1 or 2 control qubits. However, when controlled gates are decomposed, Aqcel would lose a part of opportunity to remove redundant qubit controls. More details about the decomposition are described in Appendix A. Since a multi-qubit controlled gate is decomposed into a set of Toffoli gates and its mirror for uncomputation except the central gate, the optimization of controlled operations for the set of Toffoli gates can be applied to the uncomputation part as well.

For a $n$-qubit circuit composed of $N$ multi-qubit controlled-$U$ gates, each having at most $n$ control qubits, this decomposition results in at most $nN$ controlled gates. With $nN$ gates, the cost for identifying computational basis states (Sec. 2.3) increases up to $O(n^2N^2M)$ when measuring with a quantum computer. However, the cost for removing unnecessary qubit controls improves from the above exponential scaling to $O(MnN)$. More details about the resource scaling are given in Appendix B.

After the decomposition, a $C[U]$ gate can be a single unitary $U$ gate if the probability of observing $|1\rangle$ of the control qubit is 1, or removed if the probability is 0. In all other cases, the $C[U]$ gate is kept. For a $C^2[X]$ gate, the similar control reduction can be performed with the probabilities of the four possible states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. If the probability of the state $|01\rangle$ (|10\rangle) is zero, one can eliminate the first (second) control from the $C^2[X]$ gate (see Eq. (11)). The following pseudocode is the full algorithm for redundant controlled operations removal.

**Algorithm 1: Redundant controlled operations removal**

```
for all $C[U]$ or $C^2[X]$ gate $g$ in the circuit do
execute circuit up to, but not including, $g$
if $g$ is a $C[U]$ gate then
measure the control qubit $q$ in the $Z$ basis multiple times
if $\{1\}$ is observed in the measurement results then
if $\{0\}$ is not observed in the measurement results then
turn $g$ into a $U$ gate acting on the target qubit
else if
else
eliminate $g$
end if
else
measure the control qubits $q_1q_2$ in the $Z$ basis multiple times
if $\{11\}$ is observed in the measurement re-
```

\[6\] Here the low-amplitude computational basis states mean a part of the original computational basis states composing of the ideal quantum state of interest, not those from quantum noise.
The AQCEL optimization protocol described in Sec. 2 has been deployed to the quantum parton shower (QPS) algorithm [16]. We show experimental results from a simulator and quantum hardware, and discuss the optimization performance in terms of the number of CNOT gates and the identity of the final states. The main purpose of this section is to demonstrate that AQCEL optimization works with the determination of bitstrings by a noisy intermediate scale quantum computer in polynomial resources, not to compare with other initial-state dependent optimization protocols.

### 3.1 Quantum parton shower algorithm

The QPS algorithm in Ref. [16] can start with a fermion that is either type $f_1$ or $f_2$. These fermions can radiate a scalar particle $\phi$ or not at a given showering step. The relevant parameters are the three couplings $g_1$, $g_2$, and $g_{12}$ between $f_1$ and $\phi$, $f_2$ and $\phi$, and $f_1 f_2$ ($f_1 f_2$) and $\phi$, respectively, where the antifermion is denoted by $\bar{f}$. The shower evolution process is simulated by repeating the step by $N_{\text{evol}}$ times. When $N_{\text{evol}}$ is small, only a small number of particles are simulated, hence the number of non-zero computational basis states is also small. In addition, since the circuit is designed to work with genetic initial states, a different set of computational basis states is occupied for each initial state, resulting in redundant controlled operations in the circuit.

The quantum circuit for $N_{\text{evol}} = 1$ step parton shower algorithm when the initial state is $|f_1\rangle$. Coupling constants are $g_1 = 2$ and $g_2 = g_{12} = 1$. Figure 2 shows the benchmark quantum circuit for the QPS algorithm with $N_{\text{evol}} = 1$.

### 3.2 Experimental setup

The AQCEL protocol focuses on circuit optimization at the algorithmic level, instead of at the level of a specific implementation using native gates for a particular quantum device. In addition to the initial-state dependent reduction of unnecessary controlled operations (Sec. 2.3 and 2.4), the AQCEL protocol performs sequential decompositions of multi-controlled gates as well as the removal of adjacent gate pairs and unused qubits. A high-level flowchart of the AQCEL protocol is shown in Fig. 3.

The AQCEL is implemented using IBM Qiskit
Counts

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Qiskit</th>
<th>tket</th>
<th>Qiskit + tket</th>
<th>AQCEL (CC)</th>
<th>AQCEL (CC) + Qiskit + tket</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNOT</td>
<td>100</td>
<td>150</td>
<td>200</td>
<td>250</td>
<td>50</td>
<td>260</td>
</tr>
<tr>
<td>Sx,X</td>
<td>100</td>
<td>150</td>
<td>200</td>
<td>250</td>
<td>50</td>
<td>260</td>
</tr>
</tbody>
</table>

Figure 4: Numbers of CNOT gates and single-qubit gates of the QPS circuit transpiled considering the ideal topology. For the AQCEL optimization, the bitstring determination at controlled gates is performed using classical calculation.

transpiled circuit from the input to the measurements, as implemented in Qiskit.

First, we examine the numbers of single-qubit gates and CNOT gates assuming an ideal topology before and after the AQCEL optimization alone. The determination of bitstrings at controlled gates is performed using classical calculation. Figure 4 shows the gate counts from the original circuit and the circuits optimized using either AQCEL, tket or Qiskit and their different combinations. It is seen that the AQCEL alone reduces gate counts drastically, and even more for CNOT gate when the AQCEL is combined with the tket and Qiskit.

Starting with 155 CNOT and 431 single-qubit gates, the AQCEL alone removes 114 CNOT and 388 single-qubit gates, in which 58 CNOT and 88 single-qubit gates are accounted for by the reduction of redundant qubit controls. The number of qubits is reduced from 14 to 13. The register \( n_\phi \), composed of only one qubit, is removed because it is used only for the case where the initial state is \( |\phi\rangle \). The entire AQCEL optimization takes about 0.58 seconds. The wall time is by far dominated by the elimination of redundant controlled operations (that accounts for 65% of the total time), followed by a sub-dominant contribu-

---

### 3.3 Results

Here we discuss results for the AQCEL optimization to the \( N_{\text{evol}} = 1 \) QPS circuit. The figure of merit is the numbers of single-qubit gates (\( S_X, X \)) and CNOT gates obtained by decomposing quantum gates in the circuit and the calculation time of the circuit. The calculation time is defined as the duration of the pulse schedule of the

---

\(^7\)This routine is documented in the pytket manual at https://cqcl.github.io/tket/pytket/api/backends.html. The routine for decomposition all gates into basis gates is always applied in front of tket.

\(^8\)The value of seed_transpiler is fixed to 1 in order to suppress the randomness of Qiskit transpiler.

\(^9\)\( R_z \) gates are not included because they can be implemented with no cost as virtual Z gates [70].

\(^10\)Removing redundant qubit controls reduces the cost of \( C^2[X] \) and \( C[U] \) decomposition into basis gates, resulting in the overall reduction of CNOT and single-qubit gate counts.
Counts

50
0
0
50
100
150
200
250
300
350
Counts

IBM Kawasaki Machine

<table>
<thead>
<tr>
<th>Qiskit</th>
<th>Qiskit +</th>
<th>AQCEL (CC)</th>
<th>AQCEL (QC, 0.1)</th>
<th>AQCEL (QC, dynamic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>300</td>
<td>250</td>
<td>200</td>
<td>150</td>
</tr>
</tbody>
</table>

Figure 5: Numbers of gates and the calculation times of the QPS circuit transpiled considering the topology of *ibm_kawasaki*. For the AQCEL optimization, the bitstring determination at controlled gates is performed using either classical calculation (denoted by CC) or measuring the computational basis states on quantum hardware (denoted by QC) with the static thresholds of 0.1 or the dynamic threshold of $s^\text{dyn}$.

Now we evaluate the performance of the optimizers with the transpilation considering the hardware topology of the *ibm_kawasaki*. In AQCEL, the bitstring determination is performed using classical calculation (denoted by CC) or quantum hardware (denoted by QC) with several thresholds. Figure 5 shows the results for the $N_{\text{eval}} = 1$ QPS circuit. The AQCEL reduces the CNOT gate counts more significantly than the case for ideal topology (Fig. 4). This reduction comes from less SWAP gates (each of which requires 3 CNOTs) for the AQCEL circuit because, once redundant qubit controls are removed, the amount of SWAP operations between multiple qubits is also suppressed. This results in much shorter calculation time for the AQCEL circuit. For the rest in the paper, the most efficient transpilation that combines the AQCEL, $t|\text{ket}$ and Qiskit is used for identification of the bitstrings at controlled gates and the fidelity measurement. The qubit counts are reduced from 14 to 13 under the dynamic threshold of $s^\text{dyn}$. Under the static thresholds, the qubit counts reduce to 13 for $0.005 \leq s' \leq 0.25$, but more significantly to 8 when $s' = 0.3$.

In the initial-state dependent circuit optimization, what is preserved is the equivalence of the final state, not the circuit itself. To evaluate the accuracy of the AQCEL optimization, we consider a classical fidelity between final states before and after the optimization, defined in terms of the probability distributions of the bitstrings observed in the measurement at the end of the circuits. This quantity, denoted as $F$ and referred to as just “fidelity” hereafter, is given by

$$F = \sum_{k} \sqrt{p_{k}^{\text{orig}} p_{k}^{\text{opt}}},$$

where the index $k$ runs over the bitstrings. The quantities $p_{k}^{\text{orig}}$ and $p_{k}^{\text{opt}}$ are the probabilities of observing $k$ in the original and optimized circuits, respectively.

In fact, we compute two fidelity values for each optimization method. For the ideal final state where any quantum error is not considered, the first fidelity, denoted $F_{\text{sim}}$, aims to quantify the amount of modifications to the final state introduced by the optimization procedure at the algorithmic level. To calculate the $F_{\text{sim}}$, both $p_{k}^{\text{orig}}$ and $p_{k}^{\text{opt}}$ are computed using the statevector simulation. The value of $F_{\text{sim}} = 1$ indicates that the final states are identical before and after the optimization (up to a possible phase difference on each of the qubits), while a deviation from unity gives a measure of how much the optimization has modified the final state from the ideal one.

The second fidelity value, $F_{\text{meas}}$, is computed using measurements with an actual quantum computer for $p_{k}^{\text{opt}}$. The $p_{k}^{\text{opt}}$ is estimated from the rate at which a bitstring occurs in a large number of repeated measurements. The $p_{k}^{\text{orig}}$ is computed using simulation, as for the $F_{\text{sim}}$. Even if $F_{\text{sim}}$ is 1, the presence of noise will make $F_{\text{meas}} < 1$, with the difference from unity getting larger when more gates (particularly CNOT gates) are present in the circuit. Removing CNOT gates to optimize the circuit will lower the overall effect of noise and raise the $F_{\text{meas}}$ value. However, when low-amplitude computational basis states are rejected by the threshold, more qubit controls are removed, which makes the final state different from the ideal one and decreases the $F_{\text{meas}}$ value. Thus, the $F_{\text{meas}}$ is a measure that reflects the trade-off of making the circuit shorter and changing the final state through the optimization. The measurements are performed 10000 times for each optimized circuit to obtain the $F_{\text{meas}}$ value, and the experiment is repeated 30 times with the same optimized circuit to finally obtain the average and the standard deviation of the $F_{\text{meas}}$. 

Accepted in *Quantum* 2022-09-01, click title to verify. Published under CC-BY 4.0.
values. Any error mitigation is not used in the measurements.

When the elimination of redundant qubit controls is performed based on measurements using a quantum computer with the static thresholds $s^f_{\ell}$, the threshold dependence of $F_{\text{sim}}$ and $F_{\text{meas}}$ values is shown in Fig. 6. With increasing $s^f_{\ell}$ value the $F_{\text{meas}}$ first increases, indicating the suppression of noise effects due to CNOT gate removal, but then worsens significantly at $s^f_{\ell} = 0.30$. This is understood from the behavior of the $F_{\text{sim}}$ value: the $F_{\text{sim}}$ stays close to unity up to $s^f_{\ell} = 0.25$ then decreases significantly, signaling that the optimization is too aggressive to maintain the final state from the original one if $s^f_{\ell} = 0.30$ is used. For the circuit considered here, the performance of the optimization appears to be best with $0.10 \leq s^f_{\ell} \leq 0.25$.

Shown in Fig. 7 is the $F_{\text{meas}}$ versus CNOT gate counts with the determination of bitstrings using classical calculation and hardware measurement. For the $s^{\text{dyn}}_{\ell}$, the optimized circuit turns out to be exactly same as the one obtained using classical calculation.\(^{11}\) For the AQCEL optimized circuits, the $F_{\text{meas}}$ values are $0.961 \pm 0.001$ for $s^{f}_{\ell} = 0.1$ and $0.911 \pm 0.002$ for the optimization using classical calculation. This demonstrates a clear improvement from the hardware measurement by removing qubit controls that trigger on low-amplitude computational basis states. The $F_{\text{sim}}$ value is $0.981$ for $s^{f}_{\ell} = 0.1$, meaning that the final state is modified slightly from the ideal one.

To further evaluate the accuracy of the AQCEL optimized circuit, a quantum state tomography (QST) is performed over the particle registers of six qubits in the $N_{\text{evol}} = 1$ QPS circuit (see Fig. 2). From the measurements of 30 circuits, each performed 4000 times, with Pauli $\{X, Y, Z\}$ observables using the quantum hardware, the density matrix is reconstructed for the optimized circuit and compared with one obtained from the original circuit to compute a fidelity (denoted by $F_{QST}$). In a statevector simulation, the $F_{QST}$ value is unity for both Qiskit+t$| ket \rangle$ and AQCEL circuit with the $s^{\text{dyn}}_{\ell}$, meaning that the optimization does not modify the final state including relative phases. The $F_{QST}$ values are measured to be $0.13$ and $0.41$ for Qiskit+t$| ket \rangle$ and AQCEL circuits, respectively, when performing QST on the quantum hardware. The $F_{QST}$ values are much smaller than unity due to hardware noise, but the AQCEL shows a clear improvement over the Qiskit+t$| ket \rangle$.

\(^{11}\)Although these circuits are identical, the $F_{\text{meas}}$ values are slightly different due to statistical uncertainty and the quantum noise of the actual device.
4 Discussion

4.1 Applicability of proposed optimization

The core component of the proposed circuit optimization is the identification of computational basis states with zero- or low-amplitudes using a quantum computer and the subsequent elimination of redundant controlled operations. Therefore, the AQCEL is expected to work more efficiently for quantum algorithms in which the quantum state has a small number of high-amplitude computational basis states. In other words, the AQCEL would not be effective if all the computational basis states have non-negligible amplitudes, especially when they are small in size because of thresholds applied on quantum hardware. A typical example is Quantum Phase Estimation [71] or Grover’s Algorithm, where an equal superposition state is created first by applying $H^\otimes n$ gates to the initial state $|0\rangle^\otimes n$ of the $n$-qubit system. However, even in this case, the AQCEL can efficiently produce a quantum circuit that approximates the original final state by ignoring low-amplitude computational basis states.

Another important aspect for the AQCEL optimization protocol is the resource needed to use a quantum computer for the optimization. In the NISQ era, it is worth spending the quantum computer resource in the optimization if the resulting circuit can produce higher-fidelity results than the original circuit does. In the fault-tolerant quantum computing era, using a quantum computer for the optimization may not be crucial due to its capability of correcting quantum errors during the operation. However, the initial-state dependent AQCEL optimization will be still useful for simplifying the quantum circuit, even in the fault-tolerant regime.

4.2 Further simplifications with initial-state dependent optimization

The new method proposed here for obtaining bit-strings using a quantum computer will open new possibilities in initial-state dependent quantum circuit optimization within polynomial computational resources. In this paper, we discuss the simplest example of this type of optimizations, that is, just use the information of control qubits and optimize multi-controlled qubit gates individually. There are several possibilities to extend the idea for further simplifications, e.g., using ancilla qubits, quantum gates with qutrit (3-level) states, or adding the information of target qubit in controlled gate like RPO [49]. One can optimize not only individual gates but also multiple gates as a gate set in future.

As another interesting possibility, if a circuit turns out to contain only a small number of basis states, one could represent the circuit state using fewer qubits than the original ones. Given that this approach might require a completely new computational basis, this is left for future work.

4.3 Mitigations of quantum errors

The threshold choice in AQCEL has significant impacts on $F_{\text{sim}}$ and $F_{\text{meas}}$, as seen in Figs. 6 and 7. The measurement error can be improved by adapting the unfolding technique developed in Ref. [17] and related approaches that use fewer resources [20, 72, 73, 74, 75] or further mitigate the errors [22]. A substantial contribution to the gate errors originates from CNOT gates. There are a variety of approaches to mitigate these errors, including the zero noise extrapolation with identity insertions, first proposed in Ref. [23] and generalized in Ref. [27]. The method based on the CNOT error mitigation may improve the accuracy of optimizations and the fidelity of our approach.

5 Conclusion and outlook

We have proposed a new optimization protocol, called AQCEL, for analyzing quantum circuits to remove redundant controlled operations. The AQCEL can remove unnecessary qubit controls from multi-controlled gates even when all the relevant qubits are entangled. The heart of the redundant controlled operations removal resides in the identification of zero- or low-amplitude computational basis states. In particular, this procedure can be performed through measurements using a quantum computer in polynomial time, instead of classical calculation that scales exponentially with the number of qubits. Although removing qubit controls that trigger on low-amplitude basis states will result in a circuit that produces the final state distinct from the original one, this may be a desirable feature un-
under the existence of hardware noise.

We have adopted the AQCEL optimization scheme to the quantum parton shower simulation using the ibm_kawasaki. In the experiment, the proposed scheme has shown a significant reduction of gate counts and improved the $F_{\text{meas}}$ value while retaining the accuracy of the probability distributions of the final state. For the AQCEL optimized circuits, the $F_{\text{meas}}$ values are 0.961±0.001 for the static threshold $s_f = 0.1$ and 0.911±0.002 for the optimization using classical calculation. For the dynamic threshold $s_f^{\text{dyn}}$, the optimized circuit is exactly same as the one obtained using classical calculation. The initial-state dependent optimization discussed here opens new possibilities to extend quantum circuit optimization further in future.

Acknowledgements

We acknowledge the use of IBM Quantum Services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

This study is partly carried out under the project “Optimization of HEP Quantum Algorithms” supported by the U.S.-Japan Science and Technology Cooperation Program in High Energy Physics.

CWB and BN are supported by the U.S. Department of Energy, Office of Science under contract DE-AC02-05CH11231. In particular, support comes from Quantum Information Science Enabled Discovery (QuantISED) for High Energy Physics (KA2401032).

We would like to thank Ross Duncan and Bert de Jong for useful discussions about the ZX-calculus.

A Decomposition of multi-controlled gates

When the same qubit controls are removed from $C^m[U]$ and $C^s[U] (m > s)$, Eq. (11) indicates that the condition for the removal of qubit controls is stricter for the latter because the number of $k$ is larger. This means that if we decompose a $C^m[U]$ into controlled gates with smaller number of control qubits, e.g., Toffoli and two-qubit gates, the opportunity to remove redundant qubit controls is partly lost. This suggests that removing redundant qubit controls should be applied before decomposing multi-controlled gates into basis gates like CNOT. As mentioned in Sec. 2.2 and Appendix B, the decomposition of multi-controlled gates is essential in terms of computational resources. This means that there is a trade-off between the opportunities of removing qubit controls and computational complexity.

B Computational resources for the proposed optimization scheme

The computational costs to perform the proposed optimization scheme are evaluated here. We consider a quantum circuit that contains $n$ qubits and $N$ multi-qubit controlled gates, each acting on $m$ control qubits and one target qubit.

The first step in the optimization scheme is the identification of computational basis states. If we use the classical calculation to simply track all the computational basis states whose amplitudes may be nonzero at each point of the circuit, it requires the computation of $\mathcal{O}(N2^n)$ states which grows exponentially with $n$. This method requires less computational resource than a statevector simulation, but it neglects certain rare cases where exact combinations of amplitudes lead to the elimination of redundant controlled operations. This is because a statevector simulation is applied as a classical calculation for the identification of bitstrings in AQCEL. If we measure the control qubits at each controlled gate $M$ times using a quantum computer, the total number of gate operations and measurements is given by

$$M\{m + (1 + m) + (2 + m) + \cdots + (N - 1 + m)\} = \frac{1}{2}MN(N - 1) + mMN$$

(19)

Therefore, the computational cost grows polynomially with $n$ in $\mathcal{O}(MN^2 + mMN)$.

We next consider the decision of all redundant qubit controls from a controlled gate with $m$ control qubits. Using a quantum computer that measures the $m$ control qubits $M$ times, the measured number of bitstrings is $M$. For the classical calculation, the number of basis states is $2^m$. Imagine that we choose an arbitrary combination among $2^m$ possible combinations of new qubit controls on the same controlled gate. We need to check
if all specified bitstrings in Eq. (11) for the chosen combination are not measured. The cost is $\mathcal{O}(Mm2^m)$ for one chosen combination because the size of the bitstring is $m$ and the numbers of measured and specified bitstrings are $M$ and $2^m$, respectively. Therefore, the overall computational cost for the determination of redundant qubit controls is $\mathcal{O}(Mm^42^mN)$ for $N$ multi-qubit controlled gates. The classical calculation requires $\mathcal{O}(m8^mN)$ as well.

In the AQCEL protocol, all controlled gates in the circuit are decomposed into Toffoli gates and two-qubit controlled-$U$ gates. With this decomposition, the total number of gate operations and measurement increases due to $\mathcal{O}(m)$ times more controlled gates. However, the computational cost for the redundant qubit control identification becomes polynomial in $\mathcal{O}(MmN)$ because all controlled gates have the constant number of control qubits ($m = 1, 2$). The computational cost for the identification of computational basis states still behaves polynomially in $\mathcal{O}(m^2MN^2)$ when a quantum computer is used. Given that a controlled gate has at most $n − 1$ control qubits, the total computational cost for the entire optimization sequence is $\mathcal{O}(n^2MN^2)$, when the computational basis state measurement is performed using a quantum computer.

References

[13] Souvik Das, Andrew J. Wildridge, Sachin B. Vaidya, and Andreas Jung. “Track clustering with a quantum annealer for primary vertex reconstruction at hadron col-


[52] D. Michael Miller, Robert Wille, and Rolf Drechsler. “Reducing reversible circuit cost by adding lines”. In 2010 40th IEEE Interna-
[73] Ming Gong, Ming-Cheng Chen, Yarui
