Scalable Hamiltonian programming in qubit arrays with nearest-neighbour couplings

Takahiro Tsunoda, Gaurav Bhole, Stephen A. Jones, Jonathan A. Jones, and Peter J. Leek
Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom
(Dated: March 18, 2020)

We consider the problem of selectively controlling couplings in a practical quantum processor with always-on interactions which are diagonal in the computational basis using sequences of local NOT gates. This methodology is well-known in NMR, and has been adapted for quantum information processing, but previous approaches do not scale efficiently for the general fully-connected Hamiltonian, where the time-complexity of finding solutions makes them only practical up to a few hundred qubits. Given the rapid growth in the number of qubits in cutting-edge quantum processors, it is of interest to investigate the applicability of this control scheme to much larger scale systems with realistic restrictions on connectivity. Here we present a scheme to find near time-optimal solutions that can be applied to practical engineered quantum systems with millions of qubits.

I. INTRODUCTION

In gate-based quantum computation [1], entangling interactions between qubits are generated by two-qubit logic gates, which are activated on demand. In order to turn the required two-qubit interactions on and off and program the system Hamiltonian, such gates are usually introduced by controllable external fields [2]. However, such interactions can also be generated directly when there are suitable coupling terms naturally present in the internal Hamiltonian [3, 4]. In this case, entanglement can be introduced without extra control hardware, at the cost of the complexity that comes from the fact that all the coupling terms are now “always-on”.

To overcome this problem, decoupling and selective recoupling techniques were developed in the Nuclear Magnetic Resonance (NMR) community. Just as a spin echo refocuses the coherent dynamics of a spin-1/2 nucleus [6], local NOT gates on selected qubits in a multi-qubit system reverse the evolution of corresponding terms in the internal Hamiltonian [7, 8]. By scheduling the timing and flip patterns properly, it is possible to rescale one-body and two-body terms to simulate any desired effective Hamiltonian [9]. This idea has been recently applied to replace core entanglement processes in important quantum algorithms, such as Trotterized quantum simulation [10] and the quantum Fourier transform [11], which showed a significant decrease in the total run time.

Superconducting qubits commonly use gates generated on demand but can also be tuned to generate always-on two-qubit interactions [12]. In fact always-on interactions are difficult to avoid in superconducting circuits due to their distributed microwave fields, and much effort is invested in mitigating these [13, 14]. Such always-on couplings are present in capacitively-coupled transmon qubits induced by their relatively small anharmonicity [15]. Here, we leverage such interaction expressed as a background Hamiltonian that contains single qubit interactions (offsets) \( \Omega_j \) and two qubit interactions (couplings) \( \omega_{jk} \) which are both diagonal in the computational basis,

\[
\mathcal{H}/\hbar = \sum_j \Omega_j \sigma_j^z/2 + \sum_{j<k} \omega_{jk} \sigma_j^z \sigma_k^z/4,
\]

and in principle superconducting quantum computers can be operated based on this NMR-like architecture. We have previously described methods for finding time-optimal rescaling sequences in systems of up to around 20 qubits, and near-optimal sequences in systems of around 100 qubits [16]. However, these methods do not scale to very large systems with thousands of qubits, as the computational time required to find such solutions grows rapidly with the number of qubits.

The most general form of the Hamiltonian above will have all offsets and all possible couplings present. However, couplings are generally only significant between nearby qubits: for example, superconducting qubits are often engineered in a square-lattice [17] and can have couplings limited to only nearest or next-nearest neighbour qubits if circuits are well microwave engineered [18]. Here we describe a method based on graph colouring for rapidly finding near-optimal sequences in these highly practical partially coupled systems. Initially we will assume the qubits to be arranged in a square lattice with identical couplings between nearest neighbours and no long range couplings at all, but both of these restrictions will subsequently be partly relaxed. Remarkably we find that in this partially-coupled scenario, which is very realistic for large-scale superconducting circuits, the refocusing and rescaling of always-on couplings can be efficiently programmed, requiring only linear time to design control sequences and linear number of control pulses, which widens its applicability to large-scale quantum simulation and fast closed-loop variational quantum algorithms.

II. REFOCUSING AND RESCALING

We begin with the smallest possible square lattice, containing four qubits with identical nearest neighbour couplings \( \omega_{j,j+1} \) and \( \omega_{14} \) only. Non nearest-neighbour pairs have no coupling, so \( \omega_{13} = \omega_{24} = 0 \). We can refocus all
the couplings in the system, as well as any offsets present, using the circuit of spin-echoes shown in Fig. 1(a). Here, the unitary $U$ describes the evolution of the system under the background Hamiltonian $H$ given by the propagator $U = \exp(-i\mathcal{H}t/\hbar)$, while the NOT gates are represented by $X$. As there is no coupling to be refocused between qubits 1 and 3 it is possible to apply the same pattern of NOT gates to them both, and similarly for qubits 2 and 4, so only two distinct patterns of NOT gates are required.

This network can also be extended to retain a single coupling, as shown in Fig. 1(b). In this network the same pattern of NOT gates is applied to qubits 3 and 4, and the corresponding coupling is retained, while the remaining couplings are refocused [7]. Remarkably these networks can be easily extended to retain any pattern of couplings in a square lattice system.

### III. SQUARE LATTICES

This extension works by colouring the square lattice, as described in [7]. The system can be described as a non-complete graph, with vertices corresponding to qubits and edges to couplings, with only some of the possible edges present. The graph can be coloured by assigning a colour to each vertex, and is said to be properly coloured, corresponding to complete refocusing, if no two connected vertices are the same colour. Thus in the original fully-decoupled square of four qubits we can colour qubits 1 and 3 black and qubits 2 and 4 white, while to retain a single coupling the pair of qubits involved must be assigned the same third colour, say red. To implement a colouring use patterns of NOT gates corresponding to distinct Walsh functions [7, 16] for each colour.

This colouring pattern can be tessellated across a lattice, as shown in Fig. 2 by colouring alternate qubits black and white. Here we show a four-by-four patch containing sixteen qubits, which can be embedded in a larger lattice, retaining a single coupling while refocusing all the other interactions. The required pulse sequence can be obtained by assigning black qubits $B$ to the first Walsh pattern, white qubits $W$ to the second, and the two red qubits $R$ to the third, to obtain the sequence

$$ U X_W U X_B U X_W U X_B, \tag{2} $$

where $X_W$ indicates that NOT gates are applied to the white and red qubits, and so on. Just like the sequence for the four-qubit system, this sequence requires only 4 time periods, but now requires $2q + 4$ pulses for a system of $q$ qubits. The total time required to implement the network for a $\pi/2$ evolution, corresponding to a controlled-NOT gate, is

$$ T = 1/2J, \tag{3} $$

where the nearest-neighbour couplings are of size $2\pi J$. This is the same time as is needed for an isolated coupling, as the retained coupling evolves at full strength.

### IV. SINGLE-QUBIT INTERACTIONS

The approach above will refocus all single-qubit interactions, but it is simple to modify the $X$ gates in the network to implement single-qubit rotations directly. This relies on the identity [19]

$$ \pi_{\phi_2} \pi_{\phi_1} = 2(\phi_2 - \phi_1)z, \tag{4} $$

so that applying two $\pi$ rotations around axes in the $xy$-plane separated by an angle $\delta$ is equivalent to performing a $z$-rotation through an angle $2\delta$. These rotations can be performed by modifying any pair of $X$ gates in any refocusing network, and as each qubit is controlled separately different rotations can be applied to different qubits at no cost in time or pulse count.

FIG. 1: Networks to (a) refocus all offsets and couplings in a square of four qubits with only nearest-neighbour couplings, and (b) keep only the coupling between qubits 3 and 4. Here $X$ indicates a NOT gate while $U$ indicates evolution under the background Hamiltonian. Labels on the left indicate which qubits the NOT gates are applied to.

FIG. 2: Retaining a single nearest neighbour coupling in a square lattice requires three colours: the main lattice is coloured alternately black and white while the two coupled qubits are coloured red. Grey dashed lines show nearest-neighbour couplings which have been refocused.
V. PARALLEL GATES

If it is desired to retain several different coupling interactions then this can be achieved most simply by applying such patterns back to back, changing the colouring at each stage, but it would clearly be more efficient to as far as possible perform evolutions in parallel. For simple cases this is easily performed, as depicted in Fig. 3.

In this figure, qubits which are part of the same coupling island, that is qubits which are connected either directly or indirectly by retained couplings, have been coloured the same colour. This simple approach will only be successful when, as shown here, all the couplings which could appear in an island are in fact present. (Note that the single yellow qubit forms an island on its own.) In such cases this simple colouring strategy will retain all the desired couplings while refocusing all the couplings between islands, reproducing the desired pattern in one go. As before, the total time required is simply $T = 1/2J$, the same as for an isolated coupling.

It might appear that this approach would require the number of colours to equal the number of islands, but in fact it is never necessary to use more than four, as islands which are completely disconnected (that is, islands which are not connected by couplings in the underlying Hamiltonian) can be safely coloured the same colour. By the four-colour-map theorem this will never require more than four colours $[20]$. In particular, the black and white background colours used in earlier figures are no longer required. These four colours can be assigned to four Walsh functions, requiring $T$ to be divided up into eight equal time periods. Assigning the most common colours to the two lowest Walsh functions, which both require two pulses, and the two rarer colours to the next two Walsh functions, which require four pulses, means that the total number of pulses required in a system of $q$ qubits lies between $2q$ and $3q$.

Many target patterns cannot, however, be achieved in this way, as they have one or missing couplings. Consider, for example, the pattern shown in Fig. 4, where the black couplings must be retained, and the dashed red couplings must not be retained although they connect qubits within the main island. Any colouring which implements all the black couplings that must be retained will also implement the unwanted red couplings, and so this target pattern cannot be implemented with any single colouring pattern.

VI. MULTIPLE COLOURINGS

There is, however, a simple method for achieving any target pattern using two sequential colourings, each using four colours. The first pattern assigns two colours to the odd-numbered rows and another two to the even numbered rows, thus ensuring that no vertical couplings can be retained. Along a row, the colour of the first qubit is arbitrary, but the following qubit must be the same colour if the corresponding coupling is to be retained, and the other colour if the coupling is to be refocused. The second pattern implements vertical couplings in an analogous way.

A pair of colourings which implements the target couplings in Fig. 4 is shown in Fig. 5. Any target pattern at all can be implemented in this way, in a total time $2T$. This total time will be divided into 16 equal time periods, separated by no more than 6q individual pulses.

It is useful to compare this time with timings found by our previous algorithm $[16]$ based on linear programming, which is guaranteed (when an exhaustive basis set is used) to find a true minimum-time solution. The computational complexity of this algorithm renders it impractical for large arrays, but for four-by-four arrays with nearest-neighbour couplings it is perfectly possible to use it. We have analysed a very large number of randomly chosen targets in four-by-four arrays, and in every case the optimal solution required a total time of either $T$ (in cases with no missing couplings, so a single colouring is possible) or $1.5T$ (in cases where this is not possible).

Thus the implementation time required for our colouring based networks is slightly longer than the absolute minimum required, but our colourings are far easier to design, with a computation time scaling only linearly with the total number of qubits, and so can be applied to systems with millions of qubits. The number of NOT gates

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig3.png}
\caption{Retaining a more complex pattern of couplings in a square lattice: this pattern requires four colours.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig4.png}
\caption{A target pattern of couplings which cannot be implemented with a single colouring pattern. Although the grey qubits are easily decoupled from each other and from the black island, it is impossible to retain all the desired black couplings without also retaining the undesired couplings shown as dashed red lines.}
\end{figure}
FIG. 5: A solution implementing the target couplings in Fig. 4 using two sequential colouring patterns.

required is also greatly reduced, from \( O(q^2) \) for linear programming based solutions to \( O(q) \). As an experimental implementation of a NOT gate is certain to introduce at least some error such a substantial reduction in the number of control pulses is highly desirable.

VII. NEXT-NEAREST NEIGHBOURS

Until now we have assumed that only nearest-neighbour interactions are important, with all others being too small to matter. While interactions will generally fall off rapidly with distance, this may not be rapid enough that all long-range interactions can be ignored. In particular the next-nearest neighbour interactions, across the diagonals of the square array, are likely to be significant. In the original two-colour refocusing scheme qubits connected by next-nearest neighbour interactions will be the same colour, both black or both white, and so these interactions will be retained rather than refocused, leading to a significant error in the final gate implementation.

This is not a problem with the general four-colour scheme, as the use of different colours for alternate rows and columns guarantees that diagonally connected qubits will be of different colours. Thus these networks automatically suppress any unwanted diagonal couplings, which can therefore simply be ignored. Suppressing even longer range couplings is more complicated, but can be achieved using larger numbers of colouring patterns with more colours used in each pattern.

VIII. DIFFERENT EVOLUTION TIMES

The parallel gates approach, however, contains a further important assumption: it is not sufficient to retain two different couplings if they are required to evolve for different times. While it is commonplace in quantum algorithms to make different couplings evolve through the same angle, for example to implement controlled-NOT gates, in many physical situations apparently equivalent couplings will have different strengths, and so will need different evolution times to achieve the same angle.

It might seem necessary to apply such gates using different echo sequences, but in fact they can be partly combined. Consider two couplings in the same group, where one requires evolution for a total time \( T_A \) and the second for a time \( T_B \), with \( T_B > T_A \). The naive way to achieve this is to use two different colouring patterns, one implementing the first coupling lasting for time \( T_A \), and another implementing the second coupling lasting for \( T_B \). In fact these periods can be carried out partly in parallel: during the first period, which lasts for time \( T_A \), both couplings are retained, while for the second period, which lasts for time \( T_B - T_A \), only the second coupling is retained. Thus both couplings can be carried out in a total time \( T_B \), and the method generalises for any number of distinct couplings strengths.

It follows that any pattern of couplings can be achieved in an evolution time equal to the sum of the longest evolution times required for horizontal and vertical couplings. An even simpler strict upper bound is that any set of nearest-neighbour gates can be implemented in no more than \( 2T_{\text{max}} \), where \( T_{\text{max}} = 1/2T_{\text{min}} \) is the evolution time required for the slowest possible gate.

IX. APPROXIMATING ANGLES

The approach above allows parallel gates to be implemented in time \( 2T_{\text{max}} \) even when all the coupling strengths in the array are different, but the resulting sequences will contain \( O(q) \) sets of evolution times and \( O(q^2) \) pulses. Designing such sequences also requires sorting all the evolution times into ascending order, which has computational time complexity \( O(q \log q) \), and they will also be impractical to implement experimentally, as the time differences between very similar times may be smaller than the clock resolution.

Rather than implementing a very large number of distinct evolution times precisely, it makes more sense to use a much smaller number of evolutions to approximate all the desired times, resulting in approximate evolution angles. This is most easily achieved by dividing the range of evolution times by successive powers of two, in effect encoding each evolution time as a binary number. Using \( k \) different evolution times results in \( k \)-bit precision, with an exponential increase in precision with a linear increase in the number of evolution times used. Thus using only 20 distinct delays will allow angles to be approximated
to a precision better than $10^{-6}$.

If the range of evolution times is small then it appears tempting to apply the composite pulse based techniques which have been used to correct evolution angles for isolated couplings \[21\]. Unfortunately this approach does not work well in systems with multiple couplings to the same qubit, and in particular cannot be applied to coupling networks including a chain of three or more retained couplings. It is possible to break these chains into shorter sections if composite pulse methods are to be applied, but this requires four separate colouring patterns to implement a general network.

\section{CONCLUSIONS}

The colouring techniques described here allow the efficient refocusing of interactions in engineered systems. While the resulting pulse sequences take slightly longer to implement than the absolute minimum time required, the computational time is vastly reduced, from \(O(4^q)\) for exhaustive linear programming, or \(O(q^6)\) for RROS \[10\], right down to \(O(q)\), rendering them practical in systems with thousands or even millions of qubits. The number of control pulses is also greatly reduced, from \(O(q^2)\) to \(O(q)\), thus reducing implementation errors. The method can handle unwanted next-nearest neighbour couplings, and is easily extended in a scalable way to systems with variable coupling strengths. These approximate-angle techniques can also be used to rescale couplings, rather than simply retaining or refocusing them.

Historically, superconducting qubits inherited many techniques from NMR, but as the number of qubits grows rapidly in cutting-edge processors \[17\], with individual qubit addressability, we are entering into a regime where we can address big questions on a scale that NMR cannot reach. For example, an NMR-like superconducting quantum processor could be used for simulating the concept of time-reversal in the Loschmidt echo \[22\], a longstanding research topic in statistical mechanics, where significant experimental development has taken place in NMR \[23\] since its pioneering realisation by Hahn's spin echo experiment. As the system size gets bigger, this leads on to interesting research questions such as decoherence in many-body system \[24\], quantum chaos \[23\] and the information paradox in black holes \[26\].

Although it is not yet common to use always-on couplings as a computational resource in superconducting qubits, such implementations might fit well into the context of Near-term Intermediate Scale Quantum (NISQ) applications \[27\]. For example, the native interaction can be used for hardware-efficient implementations of the variational quantum eigensolver (VQE) \[28\] and quantum approximate optimisation algorithms (QAOA) \[29\]. Since entangling gates are realized without external fields, there is less need to calibrate the control hardware, as the calibration problem can be focused only on local control. Furthermore, quantum error mitigation methods might allow slow native interactions to output useful results by extended time evolution \[30\]. Efficient refocusing and rescaling methods, such as those described here, will be essential to extend experiments to the next generation of devices.

\begin{thebibliography}{99}
\bibitem{1} A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Phys. Rev. A \textbf{52}, 3457 (1995).
\bibitem{2} B. E. Kane, Nature \textbf{393}, 133 (1998).
\bibitem{3} D. G. Cory, M. D. Price, and T. F. Havel, Physica D \textbf{120}, 82 (1998).
\bibitem{4} X. Zhou, Z.-W. Zhou, G.-C. Guo, and M. J. Feldman, Phys. Rev. Lett. \textbf{89}, 197903 (2002).
\bibitem{5} S. C. Benjamin and S. Bose, Phys. Rev. Lett. \textbf{90}, 247901 (2003).
\bibitem{6} E. L. Hahn, Phys. Rev. \textbf{80}, 580 (1950).
\bibitem{7} J. A. Jones and E. Knill, J. Magn. Reson. \textbf{141}, 322 (1999).
\bibitem{8} D. W. Leung, I. L. Chuang, F. Yamaguchi, and Y. Yamamoto, Phys. Rev. A \textbf{61}, 42310 (2000).
\bibitem{9} D. Leung, J. Mod. Opt. \textbf{49}, 1199 (2002).
\bibitem{10} A. Parra-Rodriguez, P. Lougovski, L. Lamata, E. Solano, and M. Sanz, Phys. Rev. A \textbf{101}, 022305 (2020).
\bibitem{11} A. Martin, L. Lamata, E. Solano, and M. Sanz, Phys. Rev. Research \textbf{2} (2020).
\bibitem{12} J. M. Gambetta, J. M. Chow, and M. Steffen, npj Quantum Information \textbf{3}, 2 (2017).
\bibitem{13} Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro, et al., Phys. Rev. Lett. \textbf{113}, 220502 (2014).
\bibitem{14} P. Mundada, G. Zhang, T. Hazard, and A. Houck, Phys. Rev. Applied \textbf{12}, 054023 (2019).
\bibitem{15} L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, et al., Nature \textbf{460}, 240 (2009).
\bibitem{16} G. Bhole, T. Tsumoda, P. J. Leek, and J. A. Jones, Phys. Rev. Applied \textbf{13}, 034002 (2020).
\bibitem{17} F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, et al., Nature \textbf{574}, 505 (2019).
\bibitem{18} P. A. Spring, S. Sosninna, T. Tsumoda, B. Vlastakis, and P. J. Leek, \textit{Modelling of TM modes in periodically-shorted cavities for circuit QED} (2019).
\bibitem{19} J. J. L. Morton, A. M. Tyrtyshkin, A. Ardavan, S. C. Benjamin, K. Porfyrikis, S. A. Lyon, and G. A. D. Briggs, Nature Physics \textbf{2}, 40 (2006).
\bibitem{20} K. Appel and W. Haken, Sci. Am. \textbf{237}(4), 108 (1977).
\bibitem{21} L. Xiao and J. A. Jones, Phys. Rev. A \textbf{73}, 032334 (2006).
\bibitem{22} W. Wisniaicki, Scholarpedia \textbf{7}, 11687 (2012).
\bibitem{23} J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, Phys. Rev. X \textbf{7}, 031011 (2017).
\bibitem{24} F. M. Cucchietti, D. A. R. Dalvit, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. \textbf{91}, 210403 (2003).
\end{thebibliography}
[25] P. R. Zangara and H. M. Pastawski, Physica Scripta 92, 033001 (2017).
[26] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe, Nature 567, 6165 (2019).
[27] J. Preskill, Quantum 2, 79 (2018).
[28] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature 549, 242246 (2017).
[29] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), 1411.4028.
[30] K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. 119, 180509 (2017).
[31] Y. Li and S. C. Benjamin, Phys. Rev. X 7, 021050 (2017).
[32] A. Kandala, K. Temme, A. D. Crookes, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature 567, 491495 (2019).