On the role of entanglement in qudit-based circuit compression

Xiaoqin Gao[1,2], Paul Appel[2], Nicolai Friis[3,2], Martin Ringbauer[4], and Marcus Huber[3,2]

[1]Department of Physics, University of Ottawa, Advanced Research Complex, 25 Templeton Street, K1N 6N5, Ottawa, ON, Canada

[2]Institute for Quantum Optics and Quantum Information – IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

[3]Atominstitut, Technische Universität Wien, Stadionallee 2, 1020 Vienna, Austria

[4]Universität Innsbruck, Institut für Experimentalphysik, Technikerstrasse 25, 6020 Innsbruck, Austria

Gate-based universal quantum computation is formulated in terms of two types of operations: local single-qubit gates, which are typically easily implementable, and two-qubit entangling gates, whose faithful implementation remains one of the major experimental challenges since it requires controlled interactions between individual systems. To make the most of quantum hardware it is crucial to process information in the most efficient way. One promising avenue is to use higher-dimensional systems, qudits, as the fundamental units of quantum information, in order to replace a fraction of the qubit-entangling gates with qudit-local gates. Here, we show how the complexity of multi-qubit circuits can be lowered significantly by employing qudit encodings, which we quantify by considering exemplary circuits with exactly known (multi-qubit) gate complexity. We discuss general principles for circuit compression, derive upper and lower bounds on the achievable advantage, and highlight the key role played by entanglement and the available gate set. Explicit experimental schemes for photonic as well as for trapped-ion implementations are provided and demonstrate a significant expected gain in circuit performance for both platforms.

Quantum computation is a disruptive technology that has irrevocably changed the way that computation is envisioned. It holds the potential for addressing a wide range of computational challenges [1], from factoring [2] and database search [3], to applications in quantum machine learning [4, 5]. Yet, current noisy intermediate-scale quantum (NISQ) devices [6] are still far from addressing these applications at a practically relevant scale and early demonstrations of quantum advantages remain confined to algorithms that do not yet have clearly identifiable broad applications [7, 8]. The primary obstacle for today's quantum computers, which now feature 10s to 100s of qubits [9–16] remains noise and decoherence, which limits the number of entangling operations and therefore the achievable circuit depth. Hence, for current as well as future quantum computers, efficiency at the circuit level will be key to getting the most out of these devices.

Fortunately, there is a lot of unused potential in current quantum devices, which tend to use only a small fraction of the available Hilbert space. Indeed, control over the inherently high-dimensional Hilbert space has been demonstrated in all major quantum technology platforms [17–20], motivating the exploitation of a new paradigm of quantum computing based on d-dimensional *qudits*, rather than qubits. Compared to their two-level counterpart, qudit architectures offer much richer coherence [21] and entanglement structures [22], which can be exploited for efficient quantum information processing [23–26] and improved quantum error correction [27, 28]. Since entangling operations tend to be the bottleneck in current quantum devices, the efficiency of a quantum computation, or the complexity of a quantum circuit, is traditionally measured by counting the number of entangling operations [29]. While this is an incomplete picture, it serves as a good hardware-agnostic approximation, because the optimal circuit for a given quantum operation is elusive and highly dependent on the available gate set.

Here, we investigate *qudit circuit compression*, as a way to simplify a given qubit circuit by rephrasing it as a qudit circuit. We achieve this in two steps:

First, the qubits are partitioned into groups of equal size such that the number of gates within the groups is maximized. Each group is then interpreted as a qudit, turning entangling gates into local gates and thus reducing the overall entangling gate count by a combinatorial factor for which we find lower and upper bounds using a graph-based approach. Second, by considering an extended gate set including not only qubit-entangling gates, but also genuine qudit-entangling gates, the number of entangling gates in the resulting qudit circuit can be further reduced, even saturating the combinatorial lower bound. To showcase these two effects, we study the compression of exemplary qubit circuits under different gate sets. We illustrate this gate compression with experimental details for two contemporary quantum technologies using qudits: photonic qudits encoded in orbital angular momentum and trapped ions with multiple addressable levels, showing that, already today, qubit circuits can be more efficiently compiled on qudit architectures. Aside from algorithmic improvements, reducing the number of quantum information carriers tends to make the system experimentally easier to control, leading to improved performance. We provide general design principles, as well as upper and lower bounds for the possible reduction in the number of gates, exemplified with explicit constructions for photonic systems and trapped ions.

Circuit compression. Quantum circuits are built from a sequence of gate operations. At the lowest level of abstraction, circuit compression is hence concerned with compiling an N-qubit unitary to an M-qudit architecture ($M < N$), which we call *gate compression*, see Fig. 1. The task is to encode the qubit circuit into the qudit architecture in such a way that the maximal number of entangling gates in the qubit circuit manifests as local gates in the resulting qudit circuit. Importantly, the remaining entangling gates retain their qubit-entangling structure, in terms of maximally generated entanglement entropy, when embedded in the qudit Hilbert space. Consequently, this procedure always reduces the amount of entanglement needed (irrespective of the available gate set for qudits), by compressing non-local gates into local ones. Consider the example of a four-qubit circuit, where qubits 1 and 2 are encoded in one qudit, and qubits 3 and 4 in another. Now all non-local gates between the original qubits 1 and 2 as well as between qubits 3 and 4 are local in the respective qudits. The non-local gates between qubits that are now en-

coded in different qudits remain non-local and maintain their tensor-product structure, which is why we call them embedded qubit gates.

Importantly, while embedded qubit gates still create two-level entanglement (i.e., equivalent to a two-qubit gate in terms of entanglement entropy), they are not necessarily easily implementable in a qudit architecture. To see this, consider the example of a CNOT gate $U_{\mathrm{CNOT}}^{(c,\,t)}$ applied to qubits 2 (control) and 3 (target) in the above example of a four-qubit register, taking the form $\mathbb{1}^{(1)} \otimes U_{\mathrm{CNOT}}^{(2,3)} \otimes \mathbb{1}^{(4)}$. When qubits 1 and 2 are encoded in the first qudit, and 3 and 4 in the second, the resulting embedded version of the original gate would have to be a subspace-agnostic operation (i.e., applying the same operation on both subspaces pertaining to the encoded qubits 1 and 4). Already for a single-qudit example with the canonical encoding $|0\rangle = |00\rangle, |1\rangle = |01\rangle, |2\rangle = |10\rangle, |3\rangle = |11\rangle$, we see that performing an operation of the form $\mathbb{1} \otimes U$, which acts on the second encoded qubit only requires the application of the same U to the subspaces $\{|0\rangle, |1\rangle\}$ and $\{|2\rangle, |3\rangle\}$ in order to realize the original tensor-product structure. While embedded qubit entangling gates still only create two-level entanglement, they are hence often not available natively. On the other hand, the qudit encoding enables new kinds of two-level entangling operations, which do not admit a tensor-product structure in the corresponding qubit circuit. Such gates provide new powerful tools for circuit compression, as discussed below, while also emphasizing the importance of the available gate set for qudit architectures. This again highlights the importance of considering the available gate set, rather than just qudit dimension, for efficient circuit compression.

While finding the ideal embedding of qubits into qudits is difficult, we can use weighted graphs to find a simplified representation of the circuit which in turn allows for powerful tools of graph theory to be employed. A weighted graph $G = (V, E)$ is a pair of a set of vertices $V = \{v_1, \ldots, v_n\}$ and a set of edges $E = \{e_1, \ldots, e_k\}$, where each edge $e_i = ((v_l, v_m), w_i)$ is a pair of vertices together with a weight w_i. To encode the non-local properties of the quantum circuit into the graph, we associate each vertex with a qubit and draw an edge whenever a non-local gate is present between two qubits. The weight of each edge is determined by counting the number of non-local gates between the respective qubits. This graph representation simplifies the circuit by ignoring the gate order, but it works for both qubits and
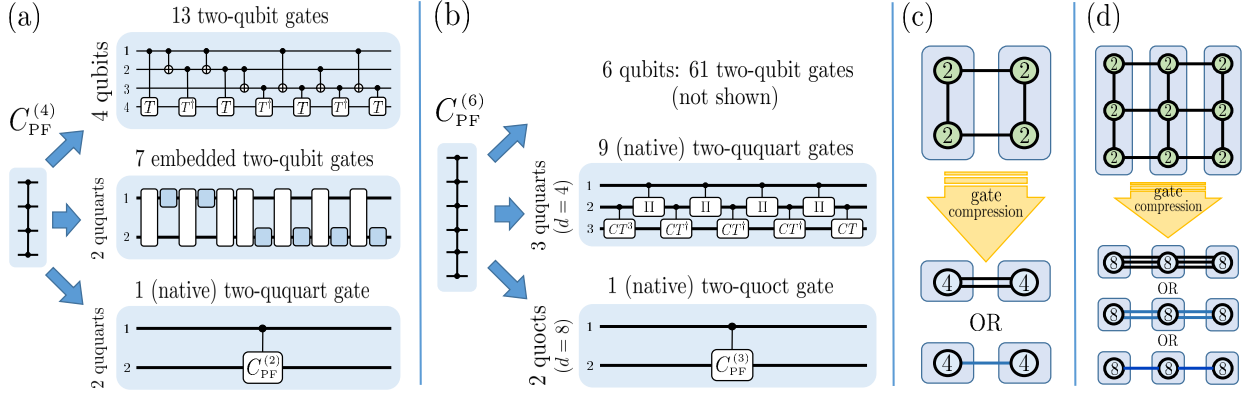
Figure 1: **Gate compression**. (a) & (b) Compression of quantum circuits of N-qubit controlled phase-flip (CPF) gates $C_{\mathrm{PF}}^{(N)}$. (a) The best known [30] decomposition of a 4-qubit CPF gate $C_{\mathrm{PF}}^{(4)}$ in terms of two-qubit gates requires 13 entangling gates, i.e., 6 CNOTs and 7 controlled T (or $T^\dagger$) gates (CT) with $T = \mathrm{diag}\{1, \exp(i\pi/4)\}$. By compressing the circuit to two qudits of dimension 4, two quarts, only 7 entangling gates are required in terms of embedded two-qubit gates, while 6 previously non-local gates become local. However, the same operation $C_{\mathrm{PF}}^{(4)}$ could also be realized by a single controlled-$C_{\mathrm{PF}}^{(2)}$ =CZ= $\mathrm{diag}\{1, 1, 1, -1\}$ gate on two ququarts, carried out if ququart 1 is in the state $|d-1\rangle = |3\rangle$ corresponding to the two-qubit state $|11\rangle$. (b) Similarly, the most efficient decomposition of $C_{\mathrm{PF}}^{(6)}$ for 6 qubits requires 61 two-qubit gates [30], whereas the compression to 3 ququarts ($d = 4$) requires (at most) 9 (native) two-ququart gates, including adjoints and powers of controlled CT gates as well as controlled Π gates, where $\Pi = \sum_{n=0}^{3} |(n-1)_{\mathrm{mod}(3)}\rangle\langle n|$ is a permutation. Further compressing to two quocts ($d = 8$), as little as 1 controlled-$C_{\mathrm{PF}}^{(3)}$ gate, conditioned on the quoct computational-basis state $|d-1\rangle = |7\rangle$ may be required. In both (a) and (b) the final decomposition of the N-qubit CPF gate into a single controlled-$C_{\mathrm{PF}}^{(N/2)}$ gate requires a non-factorizable two-level entangling gate. (c) & (d) Illustration of the effect of gate compression for the creation of a (quadratic) 2D cluster state consisting of four (c) and nine qubits (d), respectively. Cluster states are a particular type of graph state that can be schematically represented by vertices (here shown as circles carrying the local subsystem dimension d as labels, 2 for qubits, 4 for ququarts, etc.) connected by lines representing entangling gates (controlled-Z gates for qubits). By grouping (blue boxes) the qubits into pairs (c) or triples (d) the subsystem dimension becomes $d = 4$ (ququart) and $d = 8$ (quoct), respectively, but the number of non-local entangling gates can be reduced (c) from 4 two-qubit gates to either 2 embedded two-qubit gates or even 1 single-qudit gate, and (d) from 12 two-qubit gates to 6 embedded two-qubit gates, 4 two-qudit gates, or 2 two-quoct gates, as is explained below.

qudits. Using this graph representation, one can employ graph-partition algorithms [31, 32] to divide the graph into subgraphs (partitions) in a way that maximizes the number of non-local gates within each partition. Specifically, we try to find a minimal k-cut, which involves identifying a set of edges E' that, when removed, partitions the graph into k parts while minimizing the sum of weights in E'. Note that this paradigm can cover a wide range of optimization goals, including unequal-size partitions or different cost functions. In particular, we highlight the importance of considering the structure of the original circuit, as well as the available forms of qudit entanglement in the target hardware to achieve optimal encoding of qubits into qudits.

Gate set optimization. After we have reduced the *width* of the circuit, we can further improve upon the circuit *depth* by considering the extended gate set we can access for qudits. Recall that the embedded gates, resulting from circuit compression, still generate two-level entanglement and retain their original tensor-product structure. While such gates can be constructed from a universal set of qudit gates, there is much more untapped potential in the qudit Hilbert space, including gates that do not respect the tensor-product structure of the original qubit Hilbert space. By expanding the gate set, a significant reduction in circuit depth can thus be achieved, as illustrated in Fig. 1. However, the choice of gates for the expansion strongly depends on the chosen hardware platform.

Upper and lower bounds. Using the above principles, we can derive upper and lower bounds on the *compression ratio* C_d, which we define as the ratio of the number of entangling gates in the compressed circuit to the number of gates in the original circuit. To compute these bounds, we denote the weights of the original qubit graph by $\vec{\omega}$ and the weights of the qudit graph after compression by $\vec{\tilde{\omega}}$. The latter is obtained by dropping all components of the original vector that correspond to non-local gates between qubits that are

encoded in the same qudit, and which thus become local, and then assigning all edges between qubits that are encoded in different qudits to the respective qudits in the new graph (and adding the corresponding weights). The upper (worst-case) bound is then obtained by comparing the sum of weights (i.e., the 1-norm of $\vec{\omega}$) before and after compression, representing the scenario where no further gate optimization is possible. The lower bound, on the other hand, represents the scenario where gate optimization is able to fully exploit native qudit gates to realize the full qudit-entangling operation with a single gate operation. This bound is hence determined by the number of non-zero weights in $\vec{\omega}$. By normalizing these bounds using the number of gates in the initial circuit, $\|\vec{w}\|_1$, we can establish the compression-ratio bounds

$$\frac{\|\vec{\tilde{w}}\|_0}{\|\vec{w}\|_1} \leq C_d \leq \frac{\|\vec{\tilde{w}}\|_1}{\|\vec{w}\|_1}, \quad (1)$$

where, $\|\cdot\|_0$ and $\|\cdot\|_1$ refer to the l_0 (quasi-)norm and the l_1 norm, respectively. The span between the lower and upper bounds indicates potential improvements based on an appropriate gate set. In the following, we will provide examples saturating both the lower and the upper bound.

Controlled phase gates. As a key example, we will now study N-qubit controlled phase-flip gates $C_{\mathrm{PF}}^{(N)}$ (i.e., phase gates with a π phase shift). These gates are central elements in quantum computing, for example, as a key part of Grover's search algorithm [3]. It can be written as follows,

$$C_{\mathrm{PF}}^{(N)} = \mathbb{1}^{\otimes N} + |1\rangle\langle 1|^{\otimes(N-1)} \otimes (Z - \mathbb{1}), \quad (2)$$

where $\mathbb{1}$ is the identity operator and Z denotes the single-qubit Pauli-z gate. Due to its central role, the decomposition of $C_{\mathrm{PF}}^{(N)}$ is often used as a benchmark for comparing circuit decompositions and gate sets. Importantly, provably optimal decompositions into two-qubit gates and local operations are known for 3 to 8 qubits [30]. As we noted above, the optimality of a quantum circuit depends on the metric used. However, even when the metric is fixed, finding an efficient way to realize the circuit is highly non-trivial. For example, for a four-qubit controlled phase flip $C_{\mathrm{PF}}^{(4)}$, the most efficient decomposition known, believed to be optimal, requires 13 two-qubit gates [30], as shown in Fig. 1 (a). The same circuit can be realized in a qudit system by embedding two qubits each into two ququarts. There are 3 ways to achieve such a partition, and the best choice turns 6 entangling gates into local gates, leaving 7 embedded

two-qubit gates between the two ququarts. If we consider also ququart gates in our gate set, this can further reduce the non-local gate count to just a single gate between the two ququarts, which turn out to generate only two-level entanglement, as shown in Fig. 1 (a). In case $N > 4$, we can cut N qubits into k equal parts by using a $2^{N/k}$-dimensional qudit in each part and achieve similar improvements. For the six-qubit controlled phase-flip gate ($C_{\mathrm{PF}}^{(6)}$), for example, 61 two-qubit gates are needed in the best known (again, believed to be optimal) decomposition [30]. Encoding the same circuit into 3 ququarts reduces the requirement to 9 two-ququart gates, including four two-ququart controlled-shift gates and five two-ququart controlled phase gates [33], see Fig. 1 (b). Going further by encoding the gate into two qudits of dimension $d = 8$ (quocts), the gate can again be realized with a single two-level entangling qudit gate.

Graph states and the saturation of graph partitioning bounds.– Graph states are another highly relevant example. For a set of edges $E = \{e\}$, they can always be created by applying controlled-Z gates across all edges on computational-superposition product states, i.e., $|G\rangle := \prod_{e \in E} CZ_e |+\rangle^{\otimes n}$ with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. As all CZ_e commute, it directly follows that, with an unrestricted entangling power of the gate set, the lower bound $\|\vec{\tilde{w}}\|_0$ can be saturated. If the gate set is restricted to two-level entangling gates, however, the fact that a high amount of entanglement may be generated if multiple edges cut across a bipartition directly implies that the upper bound needs to be observed with $\|\vec{\tilde{w}}\|_1$ embedded qubit gates.

Consider, for example, a 4-qubit quadratic 2D cluster state as illustrated in Fig. 1 (c). Due to the symmetry of the state the partition is irrelevant and we can thus combine the first and second qubit as well as the third and fourth qubit to a qudit each without loss of generality. This reduces the number of non-local gates from 4 to 2, which saturates the upper bound $\|\vec{\tilde{w}}\|_1$. Now, we can use gate optimization to reduce the number to 1 by combining the two non-local gates between the qudits to a single genuine qudit gate, saturating the lower bound $\|\vec{\tilde{w}}\|_0$. This example demonstrates that the commutativity of the entangling gates across compressed partitions as well as the maximum amount of entanglement generated across these partitions are crucial parameters that determine how well a qubit circuit can be compressed when compiled on a qudit architecture and which gate set would be required for the qudits.

Photonic implementation.– Photonic systems are excellent candidates for gate compression. Current developments enable increased control over higher-dimensional degrees of freedom by manipulating a photon's polarization, spatial profile, temporal profile, or frequency, either separately or simultaneously. This enables the encoding of multiple bits into a single photon, as is routinely done in entanglement-based quantum communication [34–36]. Local unitary operations are easily done within a certain degree of freedom, such as spatial manipulation through multi-plane light conversion [37], frequency manipulation [38, 39], or between different degrees of freedom, for example, using polarizing beam splitters to couple path and polarization. For instance, high-dimensional Pauli X- and Z-gates, which are parts of higher-dimensional universal gate sets, have recently been implemented in a number of ways [37, 40–43]. While local gates can often be performed with near-unit efficiency and fidelity, entangling gates remain the Achilles heel of photonic information processing, as entangling two photons can be achieved probabilistically at best, leading to an exponential decrease in success probability with the number of entangling gates.

The latter aspect is where high-dimensional encodings can become particularly useful, as more information can be processed at higher fidelity and with potentially much higher success probability. Recent developments already show promising results [44, 45], including the implementations of high-dimensional multi-partite quantum gates [46] and of the SUM gate (a high-dimensional controlled-X gate) in the time and frequency degrees of freedom of photons [47]. Here, we propose a scheme for a photonic two-qudit entangling gate that grows only logarithmically in complexity with dimension and achieves a constant success probability of $1/4$ that is independent of the qudit dimension, see Appendix A for details.

Trapped-ion implementation.– Trapped ions are among the leading platforms for quantum information processing [48], where the electronic energy levels of each ion naturally provide a high-dimensional Hilbert space. Recently it was shown that such a system can be operated as a universal qudit quantum processor up to dimension 7 [19]. The qudit-gate set used in this demonstrations consisted of arbitrary local gates and two-qubit CNOT gates embedded in a qudit Hilbert space. Compared to a standard qubit CNOT, the embedded version exhibits error rates larger by roughly a factor of 2, independent of the Hilbert-space dimension. Beyond this basic gate set, it has been shown that both dominant gate mechanisms in this platform, the Mølmer-Sørensen gates [49] and light-shift gates [50], can be generalized to achieve genuine qudit entanglement. A first experimental realization of the latter demonstrated the generation of genuine qudit entanglement in a scalable fashion and with highly competitive error rates [50].

Compiling the example of Fig. 1 (a) and using state-of-the-art error rates for trapped-ion quantum processors of about 0.01 per qubit CNOT gate [48], a rough estimate suggests that the implementation of the four-qubit $C_{\mathrm{PF}}^{(4)}$ gate could achieve an error rate on the order of 0.12 with the standard two-level decomposition using 13 two-qubit entangling gates. Curiously, while it is known on the one hand that enlarging the Hilbert space locally (i.e., encoding 4 qubits into 4 qudits) can reduce the required number of gates quadratically [23], this gain is offset almost exactly in the 4-qubit case by the factor of 2 increased error rates incurred in the experimental implementation [19]. On the other hand, when two qubits each are encoded in a qudit of dimension 5 (4 computational levels and 1 auxiliary level), the required number of two-level entangling gates drops to 1 (albeit with a rotation angle equivalent to 4 qubit gates), achieving an estimated error rate of 0.04, see Appendix B for the circuit. Curiously, this is an example where it is optimal to use a two-level entangling gate in the qudit circuit, which is not an embedded qubit gate.

Conclusion.– We have explored two ways in which higher-dimensional architectures are universally beneficial to quantum computing. First, using gate compression, we can cut down on the amount of entanglement needed for a specific N-qubit circuit, and second, by exploiting the added capabilities of qudit systems in the form of richer gate sets, we can further reduce the number of non-local gates required to generate that entanglement. This highlights the key role played by entanglement and the available gate set in efficient qudit QIP. As every higher-dimensional gate can be achieved by any universal gate set applied a number of times, the advantage from larger gate sets is constant in the number of qubits. Similarly, even the most efficient partitioning of the qubit circuit leads to only a constant advantage. Such constant improvements, however, can make the difference between feasibility and failure.

Another important aspect is the potential breakdown of conventional wisdom regarding easily implementable local gates versus hard non-local gates. Once qudit dimensions get sufficiently large, the performance gap between local and non-local gates might change. Moreover, gate performance typically degrades somewhat with system size, providing another motivation for reducing the number of quantum-information carriers and making more efficient use of available resources. Finally, we emphasize that a case-by-case evaluation of the actual trade-offs is critical to finding the optimal dimensionality for a given problem and hardware platform.

Our examples provide a promising first step, showing that qudit encodings can lead to a significant reduction in gate count. Hence, this approach can greatly increase the utility of current and future quantum hardware, using only degrees of freedom that are already present in today's quantum technology. Indeed, various quantum-computing platforms have demonstrated qudit control with ever-increasing performance. Both gate compression and gate-set optimization will be central tools for making the most of the next generation of high-dimensional quantum processors, harnessing the full potential of physical quantum information carriers.

References

[1] Arkady K. Fedorov, Nicolas Gisin, Serguei M. Beloussov, and Alexander I. Lvovsky, *Quantum computing at the quantum advantage threshold: a down-to-business review,* arXiv:2203.17181 [quant-ph] (2022).

[2] Peter W. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring,* in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (IEEE, 1994) pp. 124–134.

[3] Lov K. Grover, *A fast quantum mechanical algorithm for database search,* in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, USA, 1996) pp. 212–219, arXiv:quant-ph/9605043.

[4] Vedran Dunjko and Hans J. Briegel, *Machine learning & artificial intelligence in the quantum domain: a review of recent progress,* Rep. Prog. Phys. **81**, 074001 (2018), arXiv:1709.02779.

[5] Iris Cong, Soonwon Choi, and Mikhail D. Lukin, *Quantum convolutional neural networks,* Nat. Phys. **15**, 1273 (2019), arXiv:1810.03787.

[6] John Preskill, *Quantum Computing in the NISQ era and beyond,* Quantum **2**, 79 (2018), arXiv:1801.00862.

[7] Frank Arute *et al.*, *Quantum supremacy using a programmable superconducting processor,* Nature **574**, 505 (2019), arXiv:1910.11333.

[8] Qingling Zhu *et al.*, *Quantum computational advantage via 60-qubit 24-cycle random circuit sampling,* arXiv:2109.03494 [quant-ph] (2021).

[9] Hannes Bernien, Sylvain Schwartz, Alexander Keesling, Harry Levine, Ahmed Omran, Hannes Pichler, Soonwon Choi, Alexander S. Zibrov, Manuel Endres, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin, *Probing many-body dynamics on a 51-atom quantum simulator,* Nature **551**, 579 (2017), arXiv:1707.04344.

[10] Jiehang Zhang, Guido Pagano, Paul W. Hess, Antonis Kyprianidis, Patrick Becker, Harvey Kaplan, Alexey V. Gorshkov, Zhexuan Gong,

and Christopher Monroe, *Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator,* Nature **551**, 601 (2017), arXiv:1708.01044.

[11] Johannes S. Otterbach *et al.*, *Unsupervised machine learning on a hybrid quantum computer,* arXiv:1712.05771 [quant-ph] (2017).

[12] Xi-Lin Wang, Yi-Han Luo, He-Liang Huang, Ming-Cheng Chen, Zu-En Su, Chang Liu, Chao Chen, Wei Li, Yu-Qiang Fang, Xiao Jiang, Jun Zhang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan, *18-Qubit Entanglement with Six Photons' Three Degrees of Freedom,* Phys. Rev. Lett. **120**, 260502 (2018), arXiv:1801.04043.

[13] Nicolai Friis, Oliver Marty, Christine Maier, Cornelius Hempel, Milan Holzäpfel, Petar Jurcevic, Martin B. Plenio, Marcus Huber, Christian Roos, Rainer Blatt, and Ben Lanyon, *Observation of Entangled States of a Fully Controlled 20-Qubit System,* Phys. Rev. X **8**, 021012 (2018), arXiv:1711.11092.

[14] Conor E. Bradley, Joe Randall, Mohamed H. Abobeih, Remon C. Berrevoets, Maarten J. Degen, Michiel A. Bakker, Matthew L. Markham, Daniel J. Twitchen, and Tim Hugo Taminiau, *A Ten-Qubit Solid-State Spin Register with Quantum Memory up to One Minute,* Phys. Rev. X **9**, 031045 (2019), arXiv:1905.02094.

[15] Ivan Pogorelov, Thomas Feldker, Christian D. Marciniak, Georg Jacob, Verena Podlesnic, Michael Meth, Vlad Negnevitsky, Martin Stadler, Kirill Lakhmanskiy, Rainer Blatt, Philipp Schindler, and Thomas Monz, *Compact ion-trap quantum computing demonstrator,* PRX Quantum **2**, 020343 (2021), arXiv:2101.11390.

[16] Gary J. Mooney, Gregory A. L. White, Charles D. Hill, and Lloyd C. L. Hollenberg, *Whole-Device Entanglement in a 65-Qubit Superconducting Quantum Computer,* Adv. Quantum Technol. **4**, 2100061 (2021), arXiv:2102.11521.

[17] Jianwei Wang, Stefano Paesani, Yunhong Ding, Raffaele Santagati, Paul Skrzypczyk, Alexia Salavrakos, Jordi Tura, Remigiusz Augusiak, Laura Mančinska, Davide Bacco, Damien Bonneau, Joshua W. Silverstone, Qihuang Gong, Antonio Acín, Karsten Rottwitt, Leif K. Oxenløwe, Jeremy L. O'Brien, Anthony Laing, and Mark G. Thompson, *Multidimensional quantum entanglement with large-scale in-tegrated optics,* Science **360**, 285 (2018), arXiv:1803.04449.

[18] Alexis Morvan, Vinay V. Ramasesh, Machiel S. Blok, John Mark Kreikebaum, Kevin P. O'Brien, Larry Chen, Bradley K. Mitchell, Ravi K. Naik, David I. Santiago, and Irfan Siddiqi, *Qutrit Randomized Benchmarking,* Phys. Rev. Lett. **126**, 210504 (2021), arXiv:2008.09134.

[19] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz, *A universal qudit quantum processor with trapped ions,* Nat. Phys. **18**, 1053 (2022), arXiv:2109.06903.

[20] Yulin Chi, Jieshan Huang, Zhanchuan Zhang, Jun Mao, Zinan Zhou, Xiaojiong Chen, Chonghao Zhai, Jueming Bao, Tianxiang Dai, Huihong Yuan, Ming Zhang, Daoxin Dai, Bo Tang, Yan Yang, Zhihua Li, Yunhong Ding, Leif K. Oxenløwe, Mark G. Thompson, Jeremy L. O'Brien, Yan Li, Qihuang Gong, and Jianwei Wang, *A programmable qudit-based quantum processor,* Nat. Commun. **13**, 1166 (2022), arXiv:1803.04449.

[21] Martin Ringbauer, Thomas R. Bromley, Marco Cianciaruso, Ludovico Lami, W. Y. Sarah Lau, Gerardo Adesso, Andrew G. White, Alessandro Fedrizzi, and Marco Piani, *Certification and Quantification of Multilevel Quantum Coherence,* Phys. Rev. X **8**, 041007 (2018), arXiv:1707.05282.

[22] Tristan Kraft, Christina Ritz, Nicolas Brunner, Marcus Huber, and Otfried Gühne, *Characterizing Genuine Multilevel Entanglement,* Phys. Rev. Lett. **120**, 060502 (2018), arXiv:1707.01050.

[23] Benjamin P. Lanyon, Marco Barbieri, Marcelo P. Almeida, Thomas Jennewein, Timothy C. Ralph, Kevin J. Resch, Geoff J. Pryde, Jeremy L. O'Brien, Alexei Gilchrist, and Andrew G. White, *Simplifying quantum logic using higher-dimensional Hilbert spaces,* Nat. Phys. **5**, 134 (2009), arXiv:0804.0272.

[24] Anastasiia S. Nikolaeva, Evgeniy O. Kiktenko, and Arkady K. Fedorov, *Efficient realization of quantum algorithms with qudits,* arXiv:2111.04384 [quant-ph] (2021).

[25] Evgeniy O. Kiktenko, Anastasiia S. Nikolaeva, Peng Xu, Georgy V. Shlyapnikov, and Arkady K. Fedorov, *Scalable quantum comput-*

ing with qudits on a graph, Phys. Rev. A **101**, 022304 (2020), arXiv:1909.08973.

[26] Yuchen Wang, Zixuan Hu, Barry C. Sanders, and Sabre Kais, *Qudits and High-Dimensional Quantum Computing,* Front. Phys. **8**, 479 (2020), arXiv:2008.00959.

[27] Fern H. E. Watson, Earl T. Campbell, Hussain Anwar, and Dan E. Browne, *Qudit color codes and gauge color codes in all spatial dimensions,* Phys. Rev. A **92**, 022312 (2015), arXiv:1503.08800.

[28] Earl T. Campbell, *Enhanced Fault-Tolerant Quantum Computing in d-Level Systems,* Phys. Rev. Lett. **113**, 230501 (2014), arXiv:1406.3055.

[29] Jonas Haferkamp, Philippe Faist, Naga B. T. Kothakonda, Jens Eisert, and Nicole Yunger Halpern, *Linear growth of quantum circuit complexity,* Nat. Phys. **18**, 5 (2022), arXiv:2106.05305.

[30] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter, *Elementary gates for quantum computation,* Phys. Rev. A **52**, 3457 (1995), arXiv:quant-ph/9503016.

[31] Goldschmidt Olivier and Dorit S. Hochbaum, *A Polynomial Algorithm for the K-Cut Problem for Fixed k,* Math. Oper. Res. **19**, 24 (1994).

[32] Aydin Buluç, Henning Meyerhenke, Ilya Safro, Peter Sanders, and Christian Schulz, *Recent Advances in Graph Partitioning,* in *Algorithm Engineering. Lecture Notes in Computer Science vol 9220*, edited by L. Kliemann and P. Sanders (Springer, Cham, 2016) Chap. 4, pp. 117–158, arXiv:1311.3144.

[33] Gavin K. Brennen, Stephen S. Bullock, and Dianne P. O'Leary, *Efficient circuits for exact-universal computation with qudits,* Quantum Inf. Comput. **6**, 436 (2006), arXiv:quant-ph/0509161.

[34] Alicia Sit, Frédéric Bouchard, Robert Fickler, Jérémie Gagnon-Bischoff, Hugo Larocque, Khabat Heshami, Dominique Elser, Christian Peuntinger, Kevin Günthner, Bettina Heim, Christoph Marquardt, Gerd Leuchs, Robert W. Boyd, and Ebrahim Karimi, *High-dimensional intracity quantum cryptography with structured photons,* Optica **4**, 1006 (2017), arXiv:1612.05195.

[35] Lukas Achatz, Lukas Bulla, Evelyn A. Or-

tega, Michael Bartokos, Sebastian Ecker, Martin Bohmann, Rupert Ursin, and Marcus Huber, *Simultaneous transmission of hyper-entanglement in three degrees of freedom through a multicore fiber,* npj Quantum Inf. **9**, 45 (2023), arXiv:2208.10777.

[36] Natalia Herrera Valencia, Vatshal Srivastav, Matej Pivoluska, Marcus Huber, Nicolai Friis, Will McCutcheon, and Mehul Malik, *High-Dimensional Pixel Entanglement: Efficient Generation and Certification,* Quantum **4**, 376 (2020), arXiv:2004.04994.

[37] Florian Brandt, Markus Hiekkamäki, Frédéric Bouchard, Marcus Huber, and Robert Fickler, *High-dimensional quantum gates using full-field spatial modes of photons,* Optica **7**, 98 (2020), arXiv:1907.13002.

[38] Michael Kues, Christian Reimer, Piotr Roztocki, Luis Romero Cortés, Stefania Sciara, Benjamin Wetzel, Yanbing Zhang, Alfonso Cino, Sai T. Chu, Brent E. Little, *et al.*, *On-chip generation of high-dimensional entangled quantum states and their coherent control,* Nature **546**, 622 (2017).

[39] Meritxell Cabrejo Ponce, André Luiz Marques Muniz, Marcus Huber, and Fabian Steinlechner, *High-Dimensional Entanglement for Quantum Communication in the Frequency Domain,* Laser Photonics Rev. , 2201010 (2023), arXiv:2206.00969.

[40] Ali Asadian, Paul Erker, Marcus Huber, and Claude Klöckl, *Heisenberg-Weyl observables: Bloch vectors in phase space,* Phys. Rev. A **94**, 010301(R) (2016), arXiv:1512.05640.

[41] Amin Babazadeh, Manuel Erhard, Feiran Wang, Mehul Malik, Rahman Nouroozi, Mario Krenn, and Anton Zeilinger, *High-Dimensional Single-Photon Quantum Gates: Concepts and Experiments,* Phys. Rev. Lett. **119**, 180510 (2017), arXiv:1702.07299.

[42] Xi-Lin Wang, Xin-Dong Cai, Zu-En Su, Ming-Cheng Chen, Dian Wu, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan, *Quantum teleportation of multiple degrees of freedom of a single photon,* Nature **518**, 516 (2015), arXiv:1409.7769.

[43] Xiaoqin Gao, Mario Krenn, Jaroslav Kysela, and Anton Zeilinger, *Arbitrary d-dimensional Pauli X gates of a flying qudit,* Phys. Rev. A **99**, 023825 (2019), arXiv:1811.01814.

[44] Ashok Muthukrishnan and Carlos R. Stroud Jr.,

Accepted in 〈 〉uantum 2023-09-28, click title to verify. Published under CC-BY 4.0.

8

Multivalued logic gates for quantum computation, Phys. Rev. A **62**, 052309 (2000), arXiv:quant-ph/0002033.

[45] Manuel Erhard, Robert Fickler, Mario Krenn, and Anton Zeilinger, *Twisted photons: new quantum perspectives in high dimensions,* Light Sci. Appl. **7**, 17146 (2018), arXiv:1708.06101.

[46] Xiaoqin Gao, Manuel Erhard, Anton Zeilinger, and Mario Krenn, *Computer-Inspired Concept for High-Dimensional Multipartite Quantum Gates,* Phys. Rev. Lett. **125**, 050501 (2020), arXiv:1910.05677.

[47] Poolad Imany, Jose A. Jaramillo-Villegas, Mohammed S. Alshaykh, Joseph M. Lukens, Ogaga D. Odele, Alexandria J. Moore, Daniel E. Leaird, Minghao Qi, and Andrew M. Weiner, *High-dimensional optical quantum logic in large operational spaces,* npj Quantum Inf. **5**, 1 (2019), arXiv:1805.04410.

[48] Alejandro Bermudez, Xiaosi Xu, Ramil Nigmatullin, Joe O'Gorman, Vlad Negnevitsky, Philipp Schindler, Thomas Monz, Ulrich G. Poschinger, Cornelius Hempel, Jonathan P. Home, Ferdinand Schmidt-Kaler, Michael J. Biercuk, Rainer Blatt, Simon C. Benjamin, and Markus Müller, *Assessing the Progress of Trapped-Ion Processors Towards Fault-Tolerant Quantum Computation,* Phys. Rev. X **7**, 041061 (2017), arXiv:1705.02771.

[49] Pei Jiang Low, Brendan M. White, Andrew A. Cox, Matthew L. Day, and Crystal Senko, *Practical trapped-ion protocols for universal qudit-based quantum computing,* Phys. Rev. Research **2**, 033128 (2020), arXiv:1907.08569.

[50] Pavel Hrmo, Benjamin Wilhelm, Lukas Gerster, Martin W. van Mourik, Marcus Huber, Rainer Blatt, Philipp Schindler, Thomas Monz, and Martin Ringbauer, *Native qudit entanglement in a trapped ion quantum processor,* Nat. Commun. **14**, 2242 (2023), arXiv:2206.04104.

[51] Noelia González, Gabriel Molina-Terriza, and Juan P. Torres, *How a Dove prism transforms the orbital angular momentum of a light beam,* Opt. Express **14**, 9093 (2006).

A Photonic Implementation with the orbital angular momentum

To implement quantum circuits with high-dimensional quantum gates, we propose a general experimental scheme for two-qudit CPF gates in the orbital angular momentum (OAM) of two photons. The scheme succeeds with a probability of $\frac{1}{4}$, irrespective of the encoding dimension, but requires an auxiliary two-qubit Bell state. Hence, this scheme does not overcome some of the fundamental limitations that all photonic computing suffers from but carries the potential to increase the number of gates that can be realized with a given set of resources, as well as exploring gates in one of the many potential degrees of freedom that can be harnessed in single-photons.

In order to implement an efficient photonic circuit for $C_{\mathrm{PF}}^{(N)}$ in the OAM degree-of-freedom via a qudit encoding, high-dimensional quantum gates with high success probability are required. These include single-qudit gates, which have been well developed for the OAM of a single photon, i.e., Pauli X- [37, 41, 43] and Z-gates [42], and non-local two-qudit quantum gates, which include two-qudit quantum controlled-phase gates or two-qudit quantum controlled-cyclic (permutation) gates. All of these non-local quantum gates are conditioned on the highest state $|d-1\rangle$ of the control qudit with computational basis $\{|m\rangle\}_{m=0,1,\ldots,d-1}$.

Here we are interested in performing two-qudit gates in the subspace spanned by the OAM modes $\{|m\rangle\}_{m=0,1,\ldots,d_1-1}$ and $\{|m\rangle\}_{m=0,1,\ldots,d_2-1}$ of two photons, where d_1 and d_2 are any (integer) dimensions. The main idea of our proposal for a potential experimental implementation is shown in Fig. 2 (a).
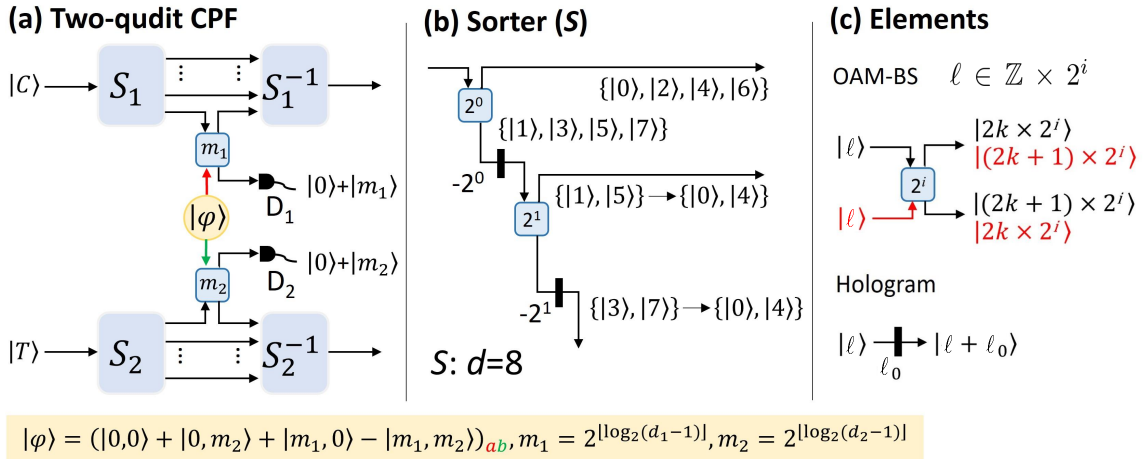

$$|\varphi\rangle = (|0,0\rangle + |0,m_2\rangle + |m_1,0\rangle - |m_1,m_2\rangle)_{ab}, m_1 = 2^{\lfloor \log_2(d_1-1)\rfloor}, m_2 = 2^{\lfloor \log_2(d_2-1)\rfloor}$$

Figure 2: **Experimental scheme for realizing a two-qudit CPF gate in the OAM degree-of-freedom of two photons**. **(a)** The setup has three main parts representing the two subsystems in question, control (upper part) and target (lower part), as well as an auxiliary state $|\varphi\rangle$ in a two-qubit subspace spanned by two OAM modes ($|0\rangle$ and $|m_1\rangle$, as well as $|0\rangle$ and $|m_2\rangle$, respectively) each of two additional photons, here shown for arbitrary local dimension d. The control and target parts each consist of single-photon input states $|C\rangle$ and $|T\rangle$ that are fed into OAM sorters S_1 and S_2, respectively, which split pairs of the different OAM modes into orthogonal spatial modes. There are $\lfloor \log_2(d_{1/2}-1)\rfloor + 1$ outputs from the sorter $S_{1/2}$. For each of these parts, the paths corresponding to the pair of OAM modes $|d_{1/2}-1-2^{\lfloor \log_2(d_{1/2}-1)\rfloor}\rangle$ and $|d_{1/2}-1\rangle$, respectively, are each combined with one of the photons of $|\varphi\rangle$ (normalization not shown) on additional OAM beam splitters (OAM-BS) $m_1 = 2^{\lfloor \log_2(d_1-1)\rfloor}$ and $m_2 = 2^{\lfloor \log_2(d_2-1)\rfloor}$ respectively. One of the output paths of each OAM-BS is then projected onto an equally weighted superposition of $|0\rangle$ and $|m_{1/2}\rangle$, while the remaining output paths of the OAM-BSs are fed back into a final inverse OAM-sorter operation $S_{1/2}^{-1}$ that recombines the different paths into a single path. **(b)** Example for OAM sorter S in dimension eight ($d = 8$) realized by a combination of two OAM-BSs [labeled by $2^i = 1$ and $2^i = 2$ with $i = 0$ and $i = 1$, respectively, as shown in **(c)**] and two holograms [for $\ell_0 = -1$ and $\ell_0 = -2$ as shown in **(c)**]. The first OAM-BS ($2^i = 1$, $i \in \mathbb{Z}^{0+}$) separates the modes $|0\rangle$, $|2\rangle$, $|4\rangle$, and $|6\rangle$ from $|1\rangle$, $|3\rangle$, $|5\rangle$, and $|7\rangle$. The first hologram ($\ell_0 = -1$) shifts the latter modes to $|0\rangle$, $|2\rangle$, $|4\rangle$, and $|6\rangle$, before the second OAM-BS ($2^i = 2$) separates $|0\rangle$ and $|4\rangle$ (formerly $|1\rangle$ and $|5\rangle$) from $|2\rangle$ and $|6\rangle$ (formerly $|3\rangle$ and $|7\rangle$). The second hologram ($\ell_0 = -2$) then maps $|2\rangle$ and $|6\rangle$ to $|0\rangle$ and $|4\rangle$, originally corresponding to $|d-1-2^{\lfloor \log_2(d-1)\rfloor}\rangle = |3\rangle$ and $|d-1\rangle = |7\rangle$. **(c)** Two basic elements: OAM beam splitters (OAM-BSs) and holograms. The label i determines the sorting property of the OAM-BS. Holograms shift the azimuthal quantum number of the OAM modes fixed amount. For details about the structure of OAM-BS see Fig. 3 (b).
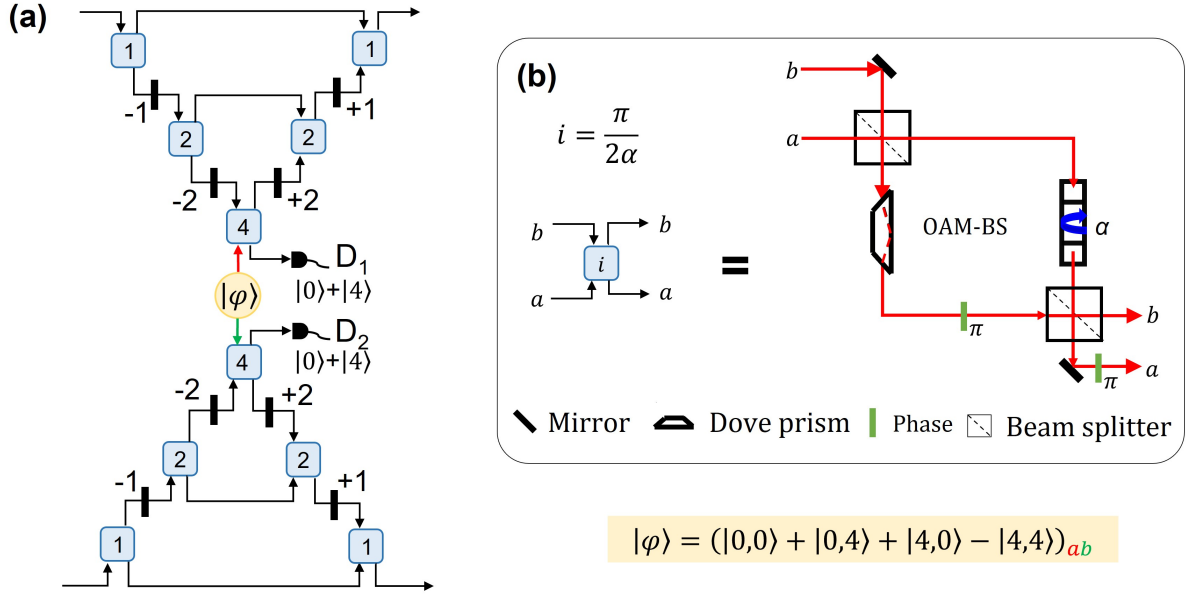
Accepted in 〈 〉uantum 2023-09-28, click title to verify. Published under CC-BY 4.0.

10

Figure 3: **Experimental scheme for two-qudit controlled phase flip in the OAM of two photons with eight dimensions**. (a) Experimental setup. 10 OAM-BSs (blue boxes) together with holograms (thick, black, vertical lines) are used in the setup. A two-qubit Bell state is necessary for facilitating the interaction between the two photons. (b) Structure of the OAM-BS. The OAM-BS works for sorting OAM modes, the sorting property is determined by i, which is related to the rotation angle α of the Dove prism.

The operation S splits modes from one path into multiple paths, while S^{-1} does the opposite, regrouping all modes into a single path. First, the input photon goes through S, which sends the highest mode $|d-1\rangle$ to its own path together with the mode $|d-1-2^{\lfloor \log_2 (d-1) \rfloor} \rangle$ by using $\lfloor \log_2 (d-1) \rfloor$ OAM-BSs and $\lfloor \log_2 (d-1) \rfloor$ holograms. Subsequently, $|d-1\rangle$ and $|d-1-2^{\lfloor \log_2 (d-1) \rfloor}\rangle$ are combined with a photon from the Bell state $|\varphi\rangle$ in an OAM-BS. Finally, the operation S^{-1} recombines all modes into one single path. The gate works if two single-photon detectors click at the same time with probability of 1/4 (due to the auxiliary Bell state, normalization not shown). As an example, we show the operation S for dimensions $d=8$ in Fig. 2 (b). The highest mode $|7\rangle$ and mode $|3\rangle$ are split into their own paths and enter into an OAM-BS together with another photon (red) from the auxiliary state. Afterward, all modes are routed into one single path via S^{-1}. The operation S consists of multiple OAM-BSs and holograms, as shown in Fig. 2 (c). An OAM-BS is an interferometer containing two Dove prisms [51], see Fig. 3. The operations S and S^{-1} features a high symmetry: S^{-1} is a mirror reflection of S, with an inversion of the hologram values.

There are two basic elements in the setup: holograms and OAM beam splitters, each of which has a finite (below 1) fidelity, meaning an advantage can only materialise in high dimensions if one needs at most a logarithmic number of these elements for a specific gate. And indeed, the number $N(d_{1/2})$ of OAM-BSs scales logarithmically with d_1 and d_2:

$$N(d_{1/2}) = 2 \times (\lfloor \log_2 (d_1 - 1) \rfloor + \lfloor \log_2 (d_2 - 1) \rfloor) + 2, \tag{3}$$

where $N(d_{1/2})$ is an upper bound obtained by studying a simple setup. For certain circuits, we know the actual number can still be lower. In the case of two eight-dimensional qudits, the complete $C_{\mathrm{PF}}^{(6)}$ can be achieved with 10 OAM-BSs, as shown in Fig. 3.

Remarkably, the probability of success is 1/4, irrespective of the dimension. Therefore, the efficiency of the circuit increases significantly when using higher-dimensional gates. It is also interesting to implement such gates in other degrees of freedom of photons, such as path.

B Trapped ion implementation

In Fig. 4, we provide a circuit for the implementation of a four-qubit CPF gate $C_{\mathrm{PF}}^{(4)}$ using either four qubits, or 2 qudits.
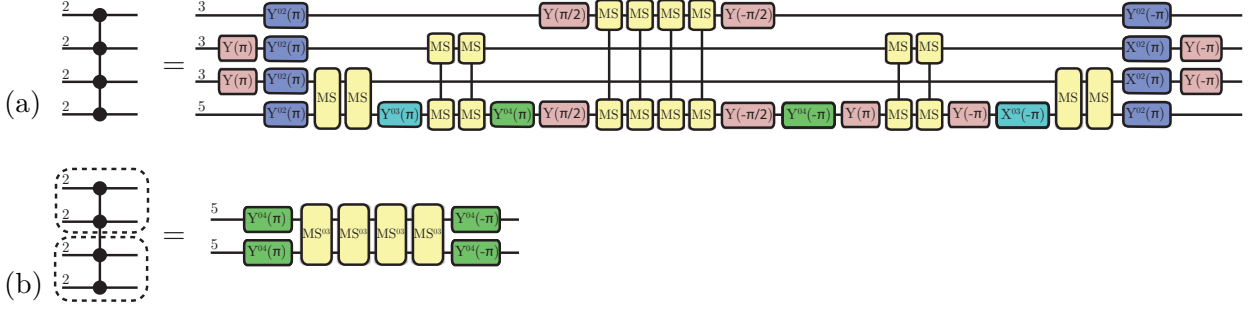

Figure 4: **Trapped-ion circuits for implementing a four-qubit CPF gate** $C_{\mathrm{PF}}^{(4)}$. Implementations are shown for (a) 4 qubits with auxiliary level, or (b) two qudits that encode 2 qubits each. In both cases only two-level entangling operations are used, where each yellow gate corresponds to a Mølmer-Sørensen (MS) gate with rotation angle $\pi/2$, which is maximally entangling for qubits. The local rotations are X or Y rotations with acting on the subspace indicated in the superscript (and color coding), where operations without superscript act on the 01-subspace of the original qubits. Note that, while using auxiliary levels enables a reduction to only 5 non-local gates, those gates require larger rotation angles and thus come with an increased experimental cost per gate. Even taking this into account the qudit-assisted circuit is expected to perform slightly better than a pure qubit circuit. In case (b) each ion occupies a 5-dimensional Hilbert space and encodes two qubits. Exploiting again the auxiliary level, only a single non-local gate is required.