Supplementary Information: Power-optimal, stabilized entangling gate between trapped-ion qubits

Reinhold Blümel,[1,2,*] Nikodem Grzesiak,[1,†] Neal Pisenti,[1] Kenneth Wright,[1] and Yunseong Nam[1,3,‡]

[1] *IonQ, College Park, MD 20740, USA*
[2] *Wesleyan University, Middletown, CT 06459, USA*
[3] *Department of Physics, University of Maryland, College Park, MD 20742, USA*

Supplementary Note 1. Resource requirement

In this section, we detail the methods used to obtain both pre- fault-tolerant (FT) and FT-regime resource requirements presented and illustrated in the main text, Fig. 1. The cases considered are (i) the water molecule ground-state energy estimation [1], (ii) Heisenberg-Hamiltonian simulation [2], (iii) quantum approximate optimization algorithm solving a maximum-cut problem [3], (iv) the quantum Fourier transform [4], (v) integer factoring [5], (vi) data-driven quantum circuit learning [6], (vii) Jellium and Hubbard-model simulation [7], and (viii) the Femoco simulation [8].

For case (i), we considered pre-FT HF+7 and HF+21 cases, where HF denotes the Hartree-Fock method detailed in [1], and 7 and 21 denote different approximation qualities. The XX gate counts for the two cases are available in Fig. 2**b** of [1].

For case (ii), we considered the Heisenberg Hamiltonian applied to spins with their connectivity specified by $(k{-}d{-}n)$ graphs, where k denotes the degree, d denotes the distance, and n denotes the number of vertices of the graph. Specifically, the graphs considered are $(3{-}5{-}70)$, $(4{-}4{-}98)$, and $(5{-}3{-}72)$. For the pre-FT cases we used CNOT gate counts reported in the pre-FT part of Table I of [2]. For the FT cases, we used T-gate counts reported in the FT part, specifically the RUS part, of the same table.

For case (iii), we considered the quantum approximate optimization algorithm in the pre-FT regime with eight stages, based on its performance compared to the well-known instance of semidefinite programming called Goemans-Williamson approximation algorithm [9]. The graphical representation of how the quantum algorithm solving the maximum cut problem performs with stage numbers $2^0, 2^1, .., 2^5$ may be found in Fig. 2 of [3]. Each stage requires $n(n-1)/2$ XX gates, as can be seen from Eq. (7) of [3].

For case (iv), we considered the approximate quantum Fourier transform [4], where all controlled-rotation gates with rotation angles less than $\pi/2^b$, $b = \log_2(n)$, where n is the number of qubits, are removed. For the pre-FT regime, one XX gate was expended per controlled-rotation gate. For the FT regime, see Table 1 of [4].

For case (v), we used the implementation presented in [5]. While an explicit resource cost is not available, an estimate is available in section A of the appendix of [10]. The implementation in [5] uses $4n^3 + O(n^2 \log(n))$ gates and $3n + 6\log(n) + O(1)$ qubits, assuming an arbitrary two-qubit gate may be implemented. For the pre-FT regime, each arbitrary two-qubit gate costs three CNOT or XX gates, as per [11]. For the FT regime, see the discussion section A of the appendix of [10], which results in $16n^3$ T gates.

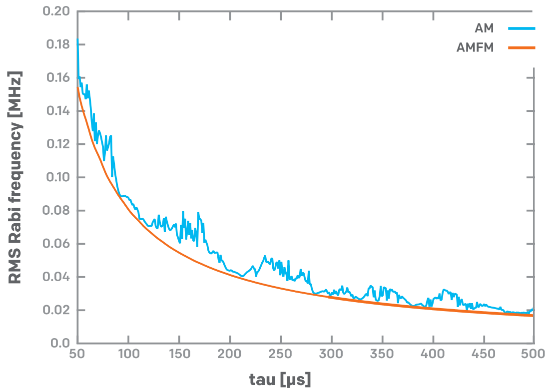
For case (vi), we largely base the resource counts on Table 1 of [6], where several sample instances of bars-and-stripes patterns are explicitly considered for n ranging from 4 to 100. The expected XX gate counts are computed assuming the all-to-all connectivity available in the trapped-ion quantum information processor (TIQIP), and we used four layers in the training circuit (see Fig. 1 of [6] for further information) that worked well for a small system with $n = 4$.

For case (vii), Tables 3 and 4 of [7] detail the FT resource-cost for several different cases.

For case (viii), see Table 1 of [8] for the Femoco simulation. We used a serial version of structure 1 with accuracy of simulation of 10^{-3} Hartree. Note this case appears to demand a fairly large amount of resources compared to other examples considered herein. This is due in part to the other examples being tailor-developed and co-designed for minimal resource requirements. Subsequent development since [8] has led to a reduction in the resource requirements by several orders of magnitude, inching closer to the cluster of points that appear in Fig. 1 of the main text. See [12] for details.

We also considered Grover's algorithm solving certain difficult instances of a Boolean satisfiability problem [13] with n variables and m clauses. Specifically, we considered "hole12", "Urq7_5", "chnl11x20', and "fpga13.12" problems, where the names were taken verbatim from Table 1 of [14]. To construct the FT circuit, we used k-control Toffoli gates to implement the Grover oracle [13], where k is the length of a clause. Specifically, we used m clean ancilla qubits to compute the satisfiability of m clauses individually, and used a m-control Toffoli gate with an additional ancilla qubit to implement the oracle. Whenever possible, we used relative Toffoli gates in [15] to reduce the T counts, while keeping track of the number of recyclable ancilla qubits in implementing the multi-control Toffoli gates. Together with a n-control Toffoli gate for the Grover diffusion operator, we

*Electronic address: blumel@ionq.co
†Electronic address: grzesiak@ionq.co
‡Electronic address: nam@ionq.co



Supplementary Figure 1: Comparison of RMS Rabi frequency required for fully entangling gates between AM and AMFM pulses. The power-optimal AM and AMFM pulses are computed in intervals of 1μs. For every gate time, the detuning of AM pulses is scanned in steps of 10Hz, and the lowest RMS Rabi frequency obtained is presented on the plot. The gates are computed for qubits $i = 1, j = 3$ in a five-ion chain, without including any stability conditions.

obtain for "hole12" 2053 qubits and $2.094 \cdot 10^{27}$ T gates, for "Urq7_5" 4627 qubits and $2.031 \cdot 10^{40}$ T gates, for "chnl11x20" 8879 qubits and $4.931 \cdot 10^{70}$ T gates, and for "fpga13.12" 2717 qubits and $1.538 \cdot 10^{39}$ T gates, where we used $\lceil \pi/4 \cdot \sqrt{2^n} \rceil$ iterations for near-optimal results. Of course it is challenging to realize on the order of $10^{27}, \ldots, 10^{70}$ quantum gates. However the presented scaling corroborates the need for efficient implementations of quantum gates in less demanding circumstances.

Supplementary Note 2. Ising gate on a trapped-ion quantum information processor

The participating ions of an Ising XX gate couple to all motional modes [16], and have to be decoupled from the motional modes at the end of the gate. The relevant equations are [17]

$$\alpha_{ip} = \int_0^\tau \Omega_i(t) \sin[\Phi_i + \psi_i(t)] e^{i\omega_p t}\, dt = 0, \quad i = 1, \ldots, N, \quad p = 1, \ldots, P, \tag{1}$$

where τ is the length of the pulse, i is the ion number, N is the total number of ions, p is the mode number, P is the total number of modes, Φ_i is the initial phase, ω_p are the motional-mode frequencies, and $\Omega_i(t)$ is the amplitude function, i.e., the time-dependent Rabi frequency. The time-dependent phases $\psi_i(t)$ in (1) are defined as

$$\psi_i(t) = \int_0^t \mu_i(t')\, dt', \tag{2}$$

where $\mu_i(t)$ is the detuning function. In order not to start the pulse abruptly, we require $\Phi_i = 0$. For ease of presentation, we also assume, from now on, that the same pulse shape acts on all N ions, such that, together with the assumption of vanishing initial phase, (1) acquires the simplified form

$$\alpha_{ip} = \int_0^\tau \Omega(t) \sin[\psi(t)] e^{i\omega_p t}\, dt = 0, \quad i = 1, \ldots, N, \quad p = 1, \ldots, P, \tag{3}$$

where

$$\psi(t) = \int_0^t \mu(t')\, dt'. \tag{4}$$

If the pulse acts simultaneously on ions i and j, the gate angle φ_{ij} of the XX gate is given by [17]

$$\varphi_{ij} = \chi_{ij} + \chi_{ji}, \tag{5}$$

where

$$\chi_{ij} = \sum_{p=1}^{P} \eta_p^i \eta_p^j \int_0^\tau dt_2 \int_0^{t_2} dt_1 \Omega(t_2)\Omega(t_1) \sin[\omega_p(t_2 - t_1)] \sin[\psi(t_2)] \sin[\psi(t_1)], \tag{6}$$

and η_p^i is the Lamb-Dicke parameter [18], which describes the coupling strength of ion number i to motional-mode number p. A maximally entangling gate is achieved for $\varphi_{ij} = \pm\pi/4$. According to (6), $\chi_{ij} = \chi_{ji}$, i.e., $\varphi_{ij} = 2\chi_{ij}$, so that a maximally entangling gate requires

$$|\chi_{ij}| = \frac{\pi}{8}. \tag{7}$$

Since both $\Omega(t)$ and $\sin[\psi(t)]$ are unknown, we combine them into one single pulse function

$$g(t) = \Omega(t) \sin[\psi(t)]. \tag{8}$$

Thus, for given motional-mode frequencies ω_p and Lamb-Dicke parameters η_p^i, our task is to find a pulse $g(t)$, which solves (3) and produces $|\chi_{ij}| = \pi/8$ with minimal power requirement. Known solution methods include amplitude-modulation techniques [17, 19], which require fixed detuning frequency μ_0, frequency-modulation techniques [20], which require a given shape of the pulse-envelope function $\Omega(t)$, and phase modulation [21]. Our approach goes beyond previously demonstrated approaches in that we modulate amplitude, frequency, and phase *simultaneously*. In addition, we use a *linear* method, which yields the optimal pulse shape directly, without any iterations or parameter searches, using exclusively linear-algebra techniques.

Supplementary Figure 1 shows a comparison between RMS Rabi frequencies, $[\langle \Omega^2(t) \rangle]^{1/2}/(2\pi)$ in MHz, required for fully entangling gates using AM and AMFM pulses. The data shows the lowest RMS Rabi frequencies obtained when scanning the gate time in steps of 1μs. For every gate time, the detuning of AM pulses is scanned in steps of 10 Hz, and the lowest RMS Rabi frequency obtained is presented on the plot. The data points were computed without any added stabilization to external sources of errors. While the required RMS Rabi frequency is lower for AMFM pulses, the stability to external errors is comparable. The execution times of the classical computations required to find the power-optimal pulses for every gate time are comparable, with 0.73 seconds for an AMFM pulse and 0.93 seconds for an AM pulse. This, however, changes drastically if higher-order moment stabilization is included, which requires the AM pulse to be represented by a large number of segments. In this case, the number of segments will approach the number of AMFM basis states and the execution time of the AMFM method will be shorter by a large factor.

Supplementary Note 3. Symmetry classes

Since $g(t)$ is a real function, the P complex equations (3) for $p = 1, \ldots, P$ are equivalent to $2P$ real equations

$$\int_0^\tau g(t)\cos(\omega_p t) \,= 0, \quad \int_0^\tau g(t)\sin(\omega_p t) \,= 0, \quad p = 1, \ldots, P. \tag{9}$$

It follows that if (9) is satisfied, any linear combination

$$h_p(t) = A_p\cos(\omega_p t) + B_p\sin(\omega_p t) \tag{10}$$

satisfies

$$\int_0^\tau g(t)h_p(t) \,= 0. \tag{11}$$

We define two special linear combinations

$$h_p^{(+)}(t) = \cos\left(\frac{\omega_p\tau}{2}\right)\cos(\omega_p t) + \sin\left(\frac{\omega_p\tau}{2}\right)\sin(\omega_p t) = \cos\left[\omega_p\left(\frac{\tau}{2} - t\right)\right] \tag{12}$$

and

$$h_p^{(-)}(t) = \sin\left(\frac{\omega_p\tau}{2}\right)\cos(\omega_p t) - \cos\left(\frac{\omega_p\tau}{2}\right)\sin(\omega_p t) = \sin\left[\omega_p\left(\frac{\tau}{2} - t\right)\right], \tag{13}$$

which satisfy

$$h_p^{(\pm)}\left(\frac{\tau}{2} - t\right) = \pm h_p^{(\pm)}\left(\frac{\tau}{2} + t\right), \tag{14}$$

i.e., $h_p^{(+)}(t)$ and $h_p^{(-)}(t)$ are even and odd functions with respect to $\tau/2$. We also define

$$g^{(\pm)}(t) = \frac{1}{2}\left[g\left(\frac{\tau}{2} + t\right) \pm g\left(\frac{\tau}{2} - t\right)\right], \tag{15}$$

i.e., the even and odd components of the pulse $g(t)$. We call $g^{(+)}(t)$ the positive-parity pulse and $g^{(-)}(t)$ the negative-parity pulse. Then the P equations

$$\int_0^\tau g^{(\pm)}(t)h_p^{(\mp)}(t)\,dt = 0, \quad p = 1, \ldots, P \tag{16}$$

are satisfied automatically, which implies that for given parity, we have to satisfy only P real, nontrivial equations

$$\int_0^\tau g^{(\pm)}(t)h_p^{(\pm)}(t)\,dt = 0, \quad p = 1, \ldots, P. \tag{17}$$

In analogy to the definition of the two parities for the pulse function $g(t)$, we may also define even and odd pulse envelope functions, $\Omega^{(\pm)}(t)$, and even and odd detuning functions, $\mu^{(\pm)}(t)$, which are even and odd functions with respect to $\tau/2$ according to

$$\Omega^{(\pm)}\left(\frac{\tau}{2} - t\right) = \pm\Omega^{(\pm)}\left(\frac{\tau}{2} + t\right), \quad \mu^{(\pm)}\left(\frac{\tau}{2} - t\right) = \pm\mu^{(\pm)}\left(\frac{\tau}{2} + t\right), \tag{18}$$

respectively. For the examples presented in this paper, we choose pulses where both the pulse-envelope function $\Omega(t)$ and the pulse-detuning function $\mu(t)$ are of positive parity. This entails that $\psi(t)$, according to (4), has odd parity with respect to $\tau/2$, so that $\sin[\psi(t)]$ is also of odd parity, resulting in a pulse function $g^{(-)}(t)$ of odd parity. Thus, to illustrate our pulse-generation method, we will in the following focus on negative-parity pulses, $g^{(-)}(t)$, constructed from a positive-parity pulse-envelope function $\Omega^{(+)}(t)$, negative-parity $\sin[\psi^{(-)}(t)]$, and positive-parity pulse-detuning function, $\mu^{(+)}(t)$. Since the pulse function is of negative parity, we expand the pulse into a Fourier-sine series according to

$$g^{(-)}(t) = \sum_{n=1}^{N_A} A_n \sin(2\pi n t/\tau), \tag{19}$$

where A_n, $n = 1, \ldots, N_A$, are real expansion amplitudes and N_A is chosen large enough to achieve convergence. The expansion (19) provides the additional benefit of switching $g^{(-)}(t)$ off continuously at $t = \tau$ without a discontinuous jump to $g = 0$ at $t = \tau$. It is straightforward to show that the expansion (19) is indeed odd with respect to $\tau/2$. The expansion (19) is complete, i.e., any pulse function $g^{(-)}(t)$ with $g^{(-)}(t = 0) = g^{(-)}(t = \tau) = 0$ can be represented this way. Expanding the entire pulse $g^{(-)}(t)$ as a whole, and not $\Omega(t)$ and $\mu(t)$ separately, is natural, since neither $\Omega(t)$ nor $\mu(t)$ are known. In fact, expansion of the entire pulse function $g(t)$ is the key idea that motivated our method of AMFM pulse construction.

Supplementary Note 4. Pulse construction

We focus in this section on computing the power-optimized pulse function $g^{(-)}(t)$ for a given set of motional-mode frequencies ω_p and Lamb-Dicke parameters η_p^i, $i = 1, \ldots, N$, $p = 1, \ldots, P$. Since in this case, according to (16), the P equations $\int_0^\tau g^{(-)}(t)h_p^{(+)}(t)dt = 0$ are automatically fulfilled, we need to fulfill, according to (17), only the set of equations

$$\int_0^\tau g^{(-)}(t)h_p^{(-)}(t)\,dt = 0, \quad p = 1, \ldots, P. \tag{20}$$

Using the expansion (19) and the explicit form (13) of $h_p^{(-)}(t)$, we obtain the following set of real, linear equations

$$\sum_{n=1}^{N_A} M_{pn} A_n = 0, \quad p = 1, \ldots, P, \tag{21}$$

where

$$M_{pn} = \int_0^\tau \sin\left(2\pi n \frac{t}{\tau}\right) \sin\left[\omega_p \left(\frac{\tau}{2} - t\right)\right] dt,$$
$$p = 1, \ldots, P, \quad n = 1, \ldots, N_A. \tag{22}$$

In matrix notation we may write (21) in the form

$$M\vec{A} = 0, \tag{23}$$

where M is the $P \times N_A$ coefficient matrix of (21) and $\vec{A}$ is the amplitude vector of length N_A. In order for (23) to have non-trivial solutions, we require $N_A > P$. In general, then, M in (23) will have rank P, and there exist $N_0 = N_A - P$ non-trivial solutions $\vec{A}^{(\alpha)}$ of (23), $\alpha = 1, \ldots, N_0$. Since $N_A > P$, the matrix M is a rectangular matrix. This suggests to multiply (23) from the left with the transpose, M^T, of M, which turns (23) into the eigenvalue problem,

$$\Gamma\vec{A} = 0, \tag{24}$$

where $\Gamma = M^T M$ is a symmetric matrix, and we are looking for the N_0 eigenvectors $\vec{A}^{(\alpha)}$ of Γ with eigenvalues 0. The N_0 nontrivial vectors $\vec{A}^{(\alpha)}$ with eigenvalues 0 span the kernel of the matrix Γ, also known as the *null space* of Γ. Numerically diagonalizing Γ, its eigenvalues typically are of the order of 10^{-12} in the null space, and several orders of magnitude larger in the complementary space. Thus, the transition from the null space to the complementary space is sharp, with eigenvalues jumping many orders of magnitude at the transition point. Therefore, the null space can be identified clearly and unambiguously. Without restriction of generality we may also assume that the null-space vectors are normalized. Since all null-space vectors $\vec{A}^{(\alpha)}$ have the common eigenvalue 0, the null space is degenerate. Thus, any linear combination of the N_0 null-space vectors $\vec{A}^{(\alpha)}$ are also null-space vectors, and we may assume that the $\vec{A}^{(\alpha)}$ form an orthonormal basis of the null space according to

$$\vec{A}^{(\alpha)\,T} \vec{A}^{(\beta)} = \delta_{\alpha\beta}, \tag{25}$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol. Our goal now is to linearly combine the orthonormal null-space vectors $\vec{A}^{(\alpha)}$ with real expansion amplitudes Λ_α to find the optimal null-space vector

$$\hat{\vec{A}} = \sum_{\alpha=1}^{N_0} \Lambda_\alpha \vec{A}^{(\alpha)} \tag{26}$$

such that

$$\hat{g}^{(-)}(t) = \sum_{n=1}^{N_A} \hat{A}_n \sin\left(2\pi n \frac{t}{\tau}\right) \tag{27}$$

is optimal in the sense that it produces $|\chi_{ij}| = \pi/8$, according to (7), and has the smallest possible norm

$$\gamma^2 = ||\hat{g}^{(-)}(t)||^2 = \frac{2}{\tau} \int_0^\tau \left[\hat{g}^{(-)}(t)\right]^2 dt = \min_{\Lambda_\alpha} \sum_{n=1}^{N_A} \hat{A}_n^2, \tag{28}$$

which entails the smallest possible average power needed to execute a maximally entangling XX gate. Using (27) with (8) and (7) in (6), we obtain

$$\frac{\pi}{8} = \left| \sum_{p=1}^{P} \eta_p^i \eta_p^j \int_0^\tau dt_2 \int_0^{t_2} dt_1 \right.$$
$$\left. \hat{g}^{(-)}(t_2)\, \hat{g}^{(-)}(t_1) \sin\left[\omega_p (t_2 - t_1)\right] \right|$$
$$= \left| \hat{\vec{A}}^T D \hat{\vec{A}} \right|, \tag{29}$$

where D is a real $N_A \times N_A$ matrix with matrix elements

$$D_{nm} = \sum_{p=1}^{P} \eta_p^i \eta_p^j \int_0^\tau dt_2 \int_0^{t_2} dt_1$$
$$\sin\left(2\pi n \frac{t_2}{\tau}\right) \sin\left[\omega_p (t_2 - t_1)\right] \sin\left(2\pi m \frac{t_1}{\tau}\right). \tag{30}$$

Since $\hat{\vec{A}}^{\,T} D \hat{\vec{A}}$ is a scalar, we can also write

$$\hat{\vec{A}}^{\,T} D \hat{\vec{A}} = \frac{1}{2}\left[\hat{\vec{A}}^{\,T} D \hat{\vec{A}} + \left(\hat{\vec{A}}^{\,T} D \hat{\vec{A}}\right)^T\right] = \hat{\vec{A}}^{\,T} S \hat{\vec{A}}, \tag{31}$$

where

$$S = \frac{1}{2}\left[D + D^T\right] \tag{32}$$

is a symmetric matrix. Using (32) and (31) in (29) we now obtain

$$\frac{\pi}{8} = \left| \vec{\Lambda}^T R \vec{\Lambda} \right|, \tag{33}$$

where $\vec{\Lambda}$ is the vector of expansion amplitudes Λ_α, $\alpha = 1, \ldots, N_0$, and R is the symmetric, reduced $N_0 \times N_0$ matrix with matrix elements

$$R_{\alpha\beta} = \vec{A}^{(\alpha)\,T} S \vec{A}^{(\beta)}, \quad \alpha, \beta = 1, \ldots, N_0. \tag{34}$$

Since R is symmetric, it can be diagonalized,

$$R\,\vec{V}^{(k)} = \lambda_k \vec{V}^{(k)}, \quad k = 1, \ldots, N_0, \tag{35}$$

where, since R is real and symmetric, the eigenvectors $\vec{V}^{(k)}$ can be assumed orthonormal. We now linearly combine the vector of expansion amplitudes $\vec{\Lambda}$ from the set of vectors $\vec{V}^{(k)}$ according to

$$\vec{\Lambda} = \sum_{k=1}^{N_0} v_k \vec{V}^{(k)}. \tag{36}$$

Supplementary Table 1: Motional-mode frequencies.

	$\omega_p/2\pi$ [MHz]
$p=1$	2.26870
$p=2$	2.33944
$p=3$	2.39955
$p=4$	2.44820
$p=5$	2.48038

Supplementary Table 2: Lamb-Dicke parameters

	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$i=1$	0.01248	0.03474	0.06091	0.07149	-0.04996
$i=2$	-0.05479	-0.07263	-0.03150	0.03406	-0.05016
$i=3$	0.08428	-0.00002	-0.05848	-0.00021	-0.05013
$i=4$	-0.05440	0.07306	-0.03098	-0.03459	-0.04991
$i=5$	0.01243	-0.03514	0.06094	-0.07163	-0.04946

According to (28), we now have to determine the expansion amplitudes v_k such that

$$\gamma^2 = \min_{v_k} \hat{\vec{A}}^T \hat{\vec{A}} = \min_{v_k} \hat{\vec{\Lambda}}^T \hat{\vec{\Lambda}} = \min_{v_k} \sum_{k=1}^{N_0} v_k^2 \tag{37}$$

under the condition

$$\frac{\pi}{8} = |\vec{\Lambda}^T R \vec{\Lambda}| = |\sum_{k=1}^{N_0} v_k^2 \lambda_k|. \tag{38}$$

Geometrically, (35) is a principal-axis transformation, $\vec{V}^{(k)}$ are the N_0 principal directions of R in the null space, (37) is a N_0-dimensional sphere of radius γ, and (38) is a N_0-dimensional conic section with principal axes $|\lambda_k|^{-1/2}$. Thus, geometrically speaking, we are looking for the smallest sphere that touches the conic section. This is obviously achieved if the sphere is inscribed in the conic section and just touches the conic section along the principal axis with the smallest length, i.e., the largest $|\lambda_k|$ Thus, our optimization problem is solved: The optimal pulse (27) is constructed with the help of the amplitudes

$$\hat{\vec{A}} = \sum_{\alpha=1}^{N_0} \Lambda_\alpha^{(k_{\max})} \vec{A}^{(\alpha)}, \tag{39}$$

where $k_{\max}$ is the index of the eigenvalue λ_k of (35) with the largest modulus $|\lambda_k|$, and

$$\vec{\Lambda}^{(k_{\max})} = v_{k_{\max}} \vec{V}^{(k_{\max})}, \tag{40}$$

where

$$v_{k_{\max}} = \left(\frac{\pi}{8|\lambda_{k_{\max}}|} \right)^{1/2}. \tag{41}$$

To illustrate the method discussed in this section, we show in the main text in Figs. 4**a** and **b** the optimal pulse $\hat{g}(t)$ obtained for $N=5$ ions and $P=5$ motional modes for mode frequencies and Lamb-Dicke parameters as shown in Tables 1 and 2, respectively. The pulse has a symmetric envelope function and is amplitude as well as frequency modulated.

Supplementary Note 5. Analytical lower bound of required peak pulse power

In this section we derive an exact, closed-form, integral-free, analytical expression for the lower bound of the minimally required pulse power needed to operate an $i \leftrightarrow j$ XX gate. To be specific, throughout this section we choose $P=N$, which is the mode in which our quantum computer is operated [22]. Generalizing the lower bound to the case $P \neq N$ is straightforward.

We define

$$G = \int_0^\tau g^2(t)\,dt = \int_0^\tau \Omega^2(t) \sin^2[\psi(t)]\,dt \le \Omega_{\max}^2 \sigma, \tag{42}$$

where we defined

$$\sigma = \int_0^\tau \sin^2[\psi(t)]\,dt. \tag{43}$$

We also define

$$\begin{aligned} D = &\sum_{pp'=1}^{N} \eta_p^i \eta_p^j \eta_{p'}^i \eta_{p'}^j \int_0^\tau dt_2 \int_0^{t_2} dt_1 \\ &\sin[\omega_p(t_2-t_1)] \sin[\omega_{p'}(t_2-t_1)] \\ = &\frac{1}{4} \sum_{p=1}^{N} (\eta_p^i \eta_p^j)^2 \left[\tau^2 - \frac{1}{\omega_p^2} \sin^2(\omega_p \tau) \right] + \\ &\sum_{p\neq p'=1}^{N} \eta_p^i \eta_p^j \eta_{p'}^i \eta_{p'}^j \left\{ \frac{1}{(\omega_p-\omega_{p'})^2} \sin^2 \left[\left(\frac{\omega_p-\omega_{p'}}{2} \right) \tau \right] \right. \\ &\left. - \frac{1}{(\omega_p+\omega_{p'})^2} \sin^2 \left[\left(\frac{\omega_p+\omega_{p'}}{2} \right) \tau \right] \right\} \le \frac{\tau^2}{4} \beta^4, \end{aligned} \tag{44}$$

where we defined

$$\beta = \left[\sum_{p=1}^{N} (\eta_p^i \eta_p^j)^2 + \sum_{p\neq p'=1}^{N} \frac{4|\eta_p^i \eta_p^j \eta_{p'}^i \eta_{p'}^j|}{(\omega_p \tau - \omega_{p'} \tau)^2} \right]^{1/4}, \tag{45}$$

which, for fixed N, is essentially a constant, which depends only weakly on τ, i.e.,

$$\beta(\tau) \sim \left[\sum_{p=1}^{N} (\eta_p^i \eta_p^j)^2 \right]^{1/4}. \tag{46}$$

For instance, for an 80 μs pulse, and the mode frequencies and η values listed in Tables 1 and 2, respectively, the first term in (45) is 2×10^{-5} while the second term is 5×10^{-8}. Therefore, in practice, the second term in (45) may be neglected. Since the Lamb-Dicke parameters η_p^j are proportional to the jth component of a unit vector [18], we have, on average, $\eta_p^j \sim 1/\sqrt{N}$, which then, because of (46), implies

$$\beta \sim 1/N^{1/4}. \tag{47}$$

With these definitions, and using the Cauchy-Schwarz inequality for integrals, we obtain:

$$\begin{aligned}
\frac{\pi}{8} &= \chi_{i,j} \\
&= \left| \sum_{p=1}^{N} \eta_p^i \eta_p^j \int_0^\tau dt_2 \int_0^{t_2} dt_1\, g(t_2) g(t_1) \sin[\omega_p(t_2 - t_1)] \right| \\
&\leq \left[\int_0^\tau dt_2 \int_0^{t_2} dt_1\, g^2(t_2) g^2(t_1) \right]^{1/2} \\
&\left\{ \int_0^\tau dt_2 \int_0^{t_2} dt_1 \left(\sum_{p=1}^{N} \eta_p^i \eta_p^j \sin[\omega_p(t_2 - t_1)] \right)^2 \right\}^{1/2} \\
&= \left[\frac{1}{2} \int_0^\tau dt_2 \int_0^\tau dt_1\, g^2(t_2) g^2(t_1) \right]^{1/2} D^{1/2} \\
&= \frac{1}{\sqrt{2}} G D^{1/2} \leq \frac{\tau\sigma}{2\sqrt{2}} \Omega_{\max}^2 \beta^2.
\end{aligned} \tag{48}$$

Using $\sin^2[\psi(t)] \leq 1$, which is valid for all arguments $\psi(t)$, the most straightforward, exact estimate for σ is

$$\sigma \leq \tau. \tag{49}$$

Using this in the inequality (48) and solving for $\Omega_{\max}$, we obtain

$$\Omega_{\max} \geq \frac{\sqrt{\pi}}{2^{3/4} \tau \beta} \tag{50}$$

or, transitioning to lab frequency,

$$f_{\max} \geq \frac{1}{2^{7/4} \sqrt{\pi} \tau \beta}. \tag{51}$$

This is the formula used to compute the analytical lower bounds of minimally required power to operate an XX gate, stated in the lower half of Table 3. The lower bound (50) [(51), respectively] is an important result. Since all the steps leading to (50) [(51), respectively] are rigorous, the lower bound (50) [(51), respectively] implies that no pulse exists, even in principle, that would require lower power than indicated by (50) [(51), respectively] to operate an XX gate. We also see that, because of (47), $\Omega_{\max}$ ($f_{\max}$, respectively) scales like $\sim N^{1/4}$.

In many cases (50) [(51), respectively] may be sharpened if lower ($\mu_{\min}$) and upper ($\mu_{\max}$) bounds for the detuning function $\mu(t)$ are available (see, e.g., the main text Fig.3**b**), i.e.,

$$\mu_{\min} \leq \mu(t) \leq \mu_{\max}, \quad t \in [0, \tau]. \tag{52}$$

We define

$$\psi_\tau = \int_0^\tau \mu(t)\, dt \leq \mu_{\max} \tau. \tag{53}$$

Supplementary Table 3: Analytical lower bounds of minimally required analytically computed peak power (lower triangle) and numerically computed peak power of optimal pulses (upper triangle) for gate combinations $i \leftrightarrow j$, mode frequencies as listed in Table 1, Lamb-Dicke parameters η as listed in Table 2, and $\tau = 300\,\mu$s. Powers quoted are in kHz. Basis size: $N_A = 1000$; $n_{\min} = 1$.

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
$i=1$	*	37.8	28.9	43.6	25.7
$i=2$	8.09	*	25.6	23.5	43.7
$i=3$	8.35	7.49	*	25.7	28.9
$i=4$	8.09	6.80	7.52	*	37.0
$i=5$	6.73	8.09	8.36	8.08	*

Since $\psi(t)$, according to (2), is defined via an integral, and since $\mu(t) > 0$ for all t, $\psi(t)$ is a monotonically increasing function of t. Therefore, in (43), we may change variables from t to ψ to obtain

$$\begin{aligned}
\sigma &= \int_{\psi_0}^{\psi_0+\psi_\tau} \sin^2(\psi) \frac{1}{\mu[t(\psi)]}\, d\psi \leq \frac{1}{\mu_{\min}} \int_{\psi_0}^{\psi_0+\psi_\tau} \sin^2(\psi)\, d\psi \\
&= \frac{1}{2\mu_{\min}} \left[\psi_\tau - \cos(2\psi_0 + \psi_\tau)\sin(\psi_\tau)\right] \leq \frac{1}{2\mu_{\min}}(\psi_\tau + 1) \\
&\leq \frac{1}{2\mu_{\min}}(\mu_{\max}\tau + 1),
\end{aligned} \tag{54}$$

where, in the last inequality, we used (53). With (54), the inequality (48) can now be stated in the form

$$\frac{\pi}{8} \leq \frac{\Omega_{\max}^2}{4\sqrt{2}\mu_{\min}} (\mu_{\max}\tau + 1)\tau\beta^2, \tag{55}$$

or, solved for $\Omega_{\max}$,

$$\Omega_{\max} \geq \frac{1}{2^{1/4}\tau\beta} \sqrt{\frac{\pi \mu_{\min}}{\mu_{\max} + 1/\tau}}. \tag{56}$$

Transitioning from angular frequency to lab frequency in Hz, we obtain

$$f_{\max} = \frac{\Omega_{\max}}{2\pi} \geq \frac{1}{2^{5/4}\sqrt{\pi}\tau\beta} \sqrt{\frac{\mu_{\min}}{\mu_{\max} + 1/\tau}}. \tag{57}$$

This is our central result. No pulse exists with a power lower than stated in (57) if χ_{ij} is determined by main text Eq. (4).

To illustrate our analytical result, we show in Table 3 a comparison between our analytical lower limit of peak pulse power and numerically obtained peak pulse powers for our sample case of $N = 5$ ions and $P = 5$ motional-mode frequencies as listed in Tables 1 and 2. We see that our analytical result is indeed lower than all numerically obtained peak pulse powers, but that both are qualitatively close.

Supplementary Note 6. Power and execution time scaling

The execution time of our linear pulse-construction algorithm is dominated by two diagonalizations, i.e., the diagonalization of the matrix $\Gamma = M^T M$ [see (24)] and the reduced matrix R [see (34)]. The dimension of Γ is $N_A \times N_A$, and the dimension of R is $(N_A - P) \times (N_A - P)$. Therefore, the execution time of our algorithm scales like $\sim N_A^3$. Since, in general, $N_A \gg P$, the execution time is dominated by N_A and depends on P only via $N_A > P$, which is needed for a nontrivial null space. Therefore, the overall scaling is dominated by N_A and the algorithm scales like $\sim N_A^3$. We confirmed the $\sim N_A^3$ scaling of our algorithm in numerous pulse-generation runs.

We also investigated the scaling of pulse power in N with up to $N = 50$ ions. For our investigation of power scaling we generated motional-mode frequencies and Lamb-Dicke parameters according to the procedure outlined in [23]. We used simulated ion positions, approximately equi-spaced with a spacing of about $5\,\mu$m and a frequency ratio of axial to radial trap frequencies of $\omega_x/\omega_r = 0.088$. We focused on operating an XX gate between ions 1 and 3. For these parameters and for $N = 50$ particles we obtained an average motional-mode frequency spacing of $\Delta f = 1.46\,$kHz. We found that our algorithm is stable only if $\tau \Delta f \approx 1$. Therefore, for our power-scaling simulations, we chose $\tau = 500\,\mu$s. The result of our power-scaling simulations in a basis of $N_A = 1000$ states is shown in the main text Fig. 2**c**. We see that the power scales approximately like $N^{1/4}$, which is consistent with the analytical power scaling (50) with (47) (gray full line in Fig. 2**c**). As pointed out in the main text, knowledge of power scaling is important since, apart from possibly damaging optical components when applying too much power, increasing power also enhances important sources of errors.

Supplementary Note 7. Power optimality requires identical pulses

In this section we show that if we do not actively stabilize the degree of entanglement χ, a gate is power optimal if ions i and j participating in a two-qubit gate are illuminated with identical laser pulses, i.e., $g_i(t) = g_j(t) = g(t)$. To show this, let $\vec{A}$ and $\vec{B}$ be the expansion amplitudes of g_i and g_j, respectively. Then, the degree of entanglement is $\chi = \vec{A}^T R \vec{B}$, where the symmetric matrix R is defined in (34). Define $P_A^2 = \vec{A}^T \vec{A}$ and $P_b^2 = \vec{B}^T \vec{B}$. Then, the task is to minimize $P^2 = P_A^2 + P_B^2$ under the constraint χ. Thus, the target function to be minimized is $F(\vec{A}, \vec{B}) = P^2 - \lambda\chi$, where λ is a Lagrangian parameter. This yields two equations:

$$\frac{\partial F}{\partial \vec{A}} = 2\vec{A} - \lambda R \vec{B} = \vec{0} \quad \Rightarrow \quad \vec{A} = \frac{1}{2}\lambda R \vec{B}, \tag{58}$$

$$\frac{\partial F}{\partial \vec{B}} = 2\vec{B} - \lambda R \vec{A} = \vec{0}, \quad \Rightarrow \quad \vec{B} = \frac{1}{2}\lambda R \vec{A}. \tag{59}$$

From (58) and (59) we obtain immediately $P_A^2 = \lambda \vec{A}^T R \vec{B}/2 = \lambda\chi/2$ and $P_B^2 = \lambda \vec{B}^T R \vec{A}/2 = \lambda \vec{A}^T R \vec{B}/2 = \lambda\chi/2$. Thus, $P_A = P_B$, i.e., for power optimality the same power must be directed at both ions. From From (58) and (59) we further obtain $\vec{A} = (\lambda R/2)(\lambda R \vec{A}/2) = \lambda^2 R^2 \vec{A}/4$ and $\vec{B} = (\lambda R/2)(\lambda R \vec{B}/2) = \lambda^2 R^2 \vec{B}/4$, i.e., $\vec{A}$ and $\vec{B}$ satisfy the same eigenvalue equation. This means that, up to normalization, $\vec{A}$ and $\vec{B}$ are the same. Together with $P_A = P_B$, we now have $g_i(t) = g_j(t) = g(t)$.

Supplementary Note 8. Stabilization against mode-frequency fluctuations

In this section we show that our linear approach lends itself naturally to a method of constructing pulses that stabilize the fidelity of the XX gate against mode drifts and mode fluctuations. Due to uncontrollable effects, such as stray electromagnetic fields, build-up of charge in the trap due to photoionization or temperature fluctuations, the frequencies of the motional modes, ω_p, will drift or fluctuate in time. Therefore, in a typical quantum-computer run, one would determine the current values of ω_p and the associated pulse $\hat{g}(t)$. However, typically over a timespan of minutes, the motional-mode frequencies ω_p will drift with typical excursions of $\Delta\omega_p/(2\pi) \approx 1\,$kHz. If we now use $\hat{g}(t)$, determined on the basis of the original mode frequencies ω_p, in the situation of the drifted modes, $\omega_p + \Delta\omega_p$, the set of equations (1) are no longer fulfilled, resulting in a reduction of the fidelity of the XX gate. A simple estimate for the infidelity increase due to the now non-zero α's in (1) is presented in [24]. According to [24], at zero temperature of the motional-mode phonons, the infidelity, $\hat{F}$, is approximately given by

$$\hat{F} = \frac{4}{5}\sum_p \left(|\alpha_{i,p}|^2 + |\alpha_{j,p}|^2\right). \tag{60}$$

This suggests stabilizing the fidelity of the quantum computer against mode drifts and fluctuations by requiring that α_{ip} be stationary up to nth order with respect to variations in ω_p. This is easily accomplished by adding the following set of equations to the set of equations (1):

$$\frac{\partial^k \alpha_{ip}}{\partial \omega_p^k} = 0 = \int_0^\tau (it)^k \Omega(t) \sin[\psi(t)] e^{i\omega_p t}\, dt,$$
$$i = 1, \ldots, N, \quad p = 1, \ldots, P, \quad k = 1, \ldots, K. \tag{61}$$

Because of the presence of the factor t^k in the integrand of (61), we call this extension of our linear approach the *moments approach*. Adding the moments equations (61) to the set (1) does not change the linearity of our method. The same techniques can be applied in solving this extended system of linear equations as was described in the main text.

Supplementary Note 9. Demodulation of pulses

The optimal pulse functions $\hat{g}(t)$ are simultaneously amplitude-, frequency-, and phase-modulated pulses. In this section we show how to demodulate the pulse $\hat{g}(t)$, i.e., how to separate $\hat{g}(t)$ into its amplitude function $\Omega(t)$ and its detuning function $\mu(t)$.

The first step of our demodulation procedure is to find the zeros ζ_j of $\hat{g}(t)$. This is numerically unproblematic, since the detuning function $\mu(t)$ is bounded away from zero, which means that degeneracies of nontrivial zeros ($\zeta_j > 0$) do not occur. In addition, in numerous simulation runs, we observed that the envelope function $\Omega(t)$ was always bounded away from zero. Therefore, in order not to complicate the discussion, we may also assume that $\Omega(t)$ does not have any zeros. Thus, all the zeros in $\hat{g}(t)$ are caused by zeros of $\sin[\psi(t)]$, i.e., $\psi(\zeta_j)$ is a multiple of π. Since no degenerate zeros occur, we have even more, namely

$$\psi(\zeta_j) = j\pi, \quad j = 0, 1, \ldots, N_z - 1, \tag{62}$$

where N_z is the total number of zeros of $\hat{g}(t)$, including the zero $\zeta_0 = 0$ at $t = 0$ and $\zeta_{N_z-1} = \tau$ at $t = \tau$. We now approximate the detuning function $\mu(t)$ as a constant between zeros of $\hat{g}(t)$, i.e.,

$$\mu(t) \approx \mu_j, \quad \zeta_{j-1} < t < \zeta_j, \quad j = 1, 2, \ldots, N_z - 1. \tag{63}$$

With (2) and (62) this entails

$$\psi(\zeta_j) - \psi(\zeta_{j-1}) = \int_{\zeta_{j-1}}^{\zeta_j} \mu(t')\, dt' = \mu_j(\zeta_j - \zeta_{j-1}) = \pi$$
$$\Longrightarrow \quad \mu_j = \frac{\pi}{\zeta_j - \zeta_{j-1}}, \quad j = 1, 2, \ldots, N_z - 1. \tag{64}$$

As an example of frequency demodulation, the main text Fig. 3**b** shows the result of the detuning function $\mu(t)$ for the pulse shown in the main text Fig. 3**a**. We see that $\mu(t)$ hovers about the middle motional mode, staying away from the strongly heating mode with the highest motional frequency. Since $g(t)$, for the example shown in Fig. 3**a**, has a dense set of $N_z = 387$ zeros, $\mu(t)$ approximated by as many piece-wise constant plateaus appears as a smooth function on the scale of Fig. 3**b**.

We now turn to extracting the pulse envelope function $\Omega(t)$ from $\hat{g}(t)$. Differentiating main text Eq. (5) and evaluating the result at the zeros ζ_j of $\hat{g}(t)$ yields

$$\begin{aligned}\hat{g}'(\zeta_j) &= \Omega'(\zeta_j)\sin[\psi(\zeta_j)] + \Omega(\zeta_j)\cos[\psi(\zeta_j)]\psi'(\zeta_j) \\ &= (-1)^j \Omega(\zeta_j)\mu(\zeta_j),\end{aligned} \tag{65}$$

where we used (4) and (62). This equation can be solved for $\Omega(\zeta_j)$ with the result

$$\Omega(\zeta_j) = (-1)^j \sigma \frac{\hat{g}'(\zeta_j)}{\mu(\zeta_j)}, \quad j = 1, \ldots, N_z - 1, \tag{66}$$

where we inserted the factor $\sigma = -\hat{g}'(\zeta_1)/|\hat{g}'(\zeta_1)|$, which ensures that $\Omega(t)$ is "right-side up", i.e., if it does not change sign, $\Omega(t) > 0$ for all t. Since $\hat{g}(t)$, according to (27), is represented by a Fourier series, it is trivial to obtain

$$\hat{g}'^{(-)}(t) = \frac{2\pi}{\tau} \sum_{n=1}^{N_A} n\hat{A}_n \cos\left(2\pi n \frac{t}{\tau}\right) \tag{67}$$

and thus $\hat{g}'^{(-)}(\zeta_j)$. The values of the detuning function $\mu(\zeta_j)$ may be obtained in several ways. We may use spline interpolation of the data set of values μ_j as defined in (64), or, as we found, with sufficient accuracy, simply use (i) $\mu(\zeta_j) = \mu_j$, (ii) $\mu(\zeta_j) = \mu_{j+1}$, or (iii) $\mu(\zeta_j) = (\mu_{j+1} + \mu_j)/2$. We used method (i) to obtain the pulse envelope function $\Omega(t)/(2\pi)$ (heavy orange line in the main text Fig. 3**a**) of the pulse $\hat{g}^{(-)}(t)$, shown as the thin green line in the main text Fig. 3**a**. The main text Fig. 3**a** shows that our amplitude demodulation technique presented above works very well and accurately extracts the envelope function.

At this point we may wonder how well the exact pulse $\hat{g}^{(-)}(t)$ is approximated by the pulse $\tilde{g}^{(-)}(t)$, i.e., the pulse reconstructed via (8) from the amplitude and detuning functions obtained by demodulating $\hat{g}^{(-)}(t)$ according to the above procedures. Therefore, to get a first impression of the accuracy of our pulse demodulation method, we compute

$$\Delta g^2 = \frac{1}{\tau}\int_0^\tau \left[\hat{g}^{(-)}(t) - \tilde{g}^{(-)}(t)\right]^2 dt, \tag{68}$$

where, for $j = 1, 2, \ldots, N_z - 1$,

$$\tilde{g}^{(-)}(t) = \Omega_j \sin[\psi_{j-1} + \mu_j(t - \zeta_{j-1})], \quad \zeta_{j-1} \le t < \zeta_j, \tag{69}$$

$$\Omega_j = (-1)^j \frac{\hat{g}'(\zeta_j)}{\mu_j}, \tag{70}$$

and

$$\psi_j = \psi_{j-1} + \mu_j(\zeta_j - \zeta_{j-1}). \tag{71}$$

Notice that Ω_j in (70) does not contain the factor σ as in (66), since this time we do not need the "right-side up" pulse, but the pulse that has the same sign of the amplitude as $\hat{g}^{(-)}(t)$. For the example shown in the main text Fig. 3**a**, we obtain $\Delta g^2 = 1.3 \times 10^{-5}$. Hence, the pulse $\tilde{g}^{(-)}(t)$ reconstructed from the demodulated pulse $\hat{g}^{(-)}(t)$ is sufficiently accurate to guarantee high-fidelity gates.

Supplementary Note 10. Stabilization against pulse-timing errors

Once τ and the mode frequencies ω_p are given, our method determines the amplitudes A_n, which are then fixed when implementing the pulse $g(t)$. However, the experimental clock may run fast or slow, which results in pulse-timing errors. To stabilize against pulse-timing errors of this nature, we require, for all $l = 0, \ldots, L$ and

$p = 1, \ldots, N$:

$$\begin{aligned}
&\frac{\partial^l}{\partial \tau^l} \int_0^\tau g(t) e^{i\omega_p t}\, dt \\
&= \frac{\partial}{\partial \tau^l} \int_0^\tau \sum_n A_n \sin(2\pi n t/\tau) e^{i\omega_p t}\, dt \\
&= \sum_n A_n \left[\frac{\partial}{\partial \tau^l} \int_0^\tau \sin(2\pi n t/\tau) e^{i\omega_p t}\, dt \right] \\
&= \sum_n A_n Q_{np}^{(l)}. \qquad (72)
\end{aligned}$$

Since $L = 0$, according to (3) of the main text, is already satisfied, (72) represents LN linear equations that may be added to the $N(K+1)$ linear equation of motional-mode stabilization. Apart from providing explicit formulas for stabilization against pulse-timing errors, this also provides a template for stabilizing against other types of errors and parameter fluctuations, and shows that it is straightforward to extend our method by any number of such linear constraints, as long as the dimension of the available null-space is not exhausted.

Supplementary Note 11. Implementation details

In this section, we present the pulse-level implementation details. Supplementary Figure 2 shows the amplitude and frequency profiles of pulses, represented according to equation (5) in the main text, with moment stabilization orders $K = 1, 4, 7$, used to implement the XX gates on our 5-qubit, 7-ion TIQIP. The motional-mode frequencies for each of the experiments are reported in Tables 4 and 5. The two-qubit gates were performed on qubits 3 and 4, with indexing starting at 0.

We note that the even-parity population, used in main text Fig. 5 to demonstrate stabilization against mode-frequency drift, is not a complete characterization of the resulting quantum gate in terms of fidelity. However, the even-parity population is the most relevant metric to use when evaluating the general experimental performance of our methods. The formalism for stabilizing phase space closure at the end of the gate with respect to gate frequency offset, relevant to Fig. 5, aims to minimize $|\alpha_p|$ in main text Eq. (3), which is most closely related to the odd-parity population, i.e., 1 minus the even-parity population. Moreover, fidelity contains within itself other sources of error accumulated during the parity contrast measurement (e.g., intensity noise differential, phase stability, laser-beam steering, overlap errors of the lasers with the positions of the trapped ions, etc.), which would mask the effect of stabilization against motional-mode drift. Thus, we decided to present the even-parity population as the most relevant quantity of interest to our work.

Mode Frequencies (MHz)
2.692
2.728
2.765
2.801
2.834
2.866
2.877

Supplementary Table 4: Mode frequencies of the motional modes of our 7-ion chain for the $K = 1$ and $K = 7$ gates shown in Supplementary Fig. 2.

Mode Frequencies (MHz)
2.690
2.726
2.763
2.799
2.832
2.864
2.876

Supplementary Table 5: Mode frequencies of the motional modes of our 7-ion chain for the $K = 4$ gate shown in Supplementary Fig. 2.

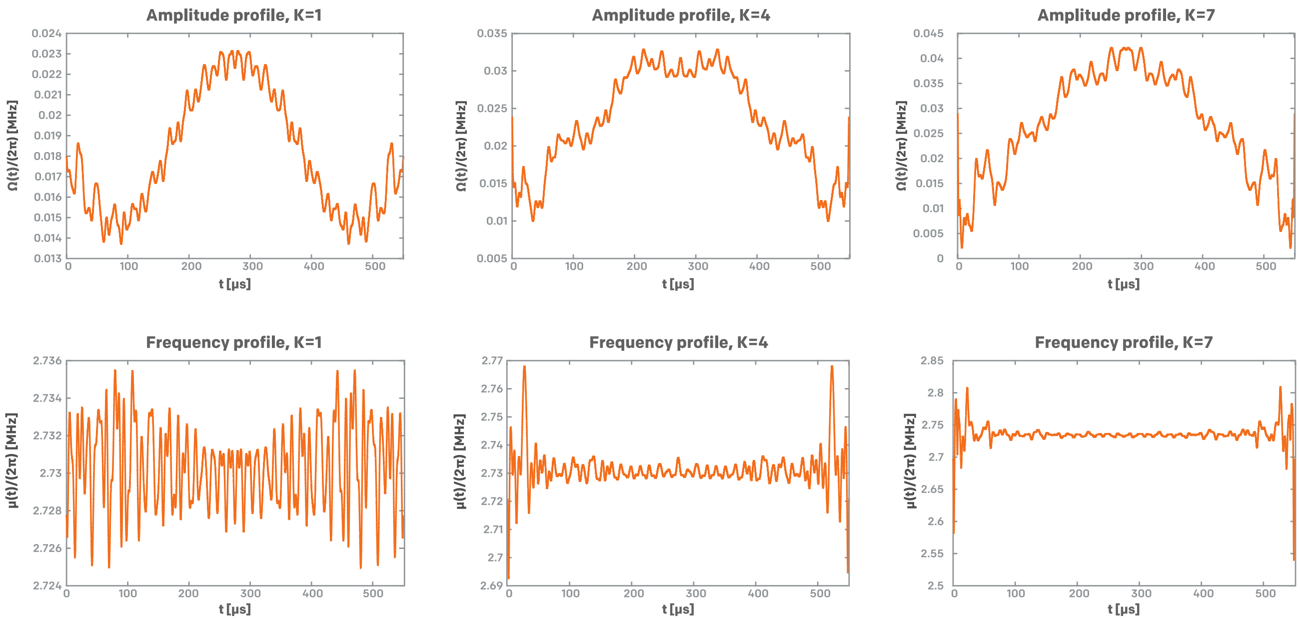
Supplementary Note 12. Fixed-detuning step pulses

Possibly the most widely studied type of fixed-detuning pulses are segmented step pulses [17]. According to this method, the detuning function $\mu(t)$ is set to a constant, i.e., $\mu(t) = \mu_0 = \text{const}$ for $0 \leq t \leq \tau$, and the pulse interval $[0, \tau]$ is broken up into $N_{\text{seg}} > P$ equi-spaced intervals $[t_{j-1}, t_j]$, $t_0 = 0$, $t_j = j\Delta t$, $\Delta t = \tau/N_{\text{seg}}$, $j = 1, \ldots, N_{\text{seg}}$, in which the pulse amplitude is set to a constant, i.e.,

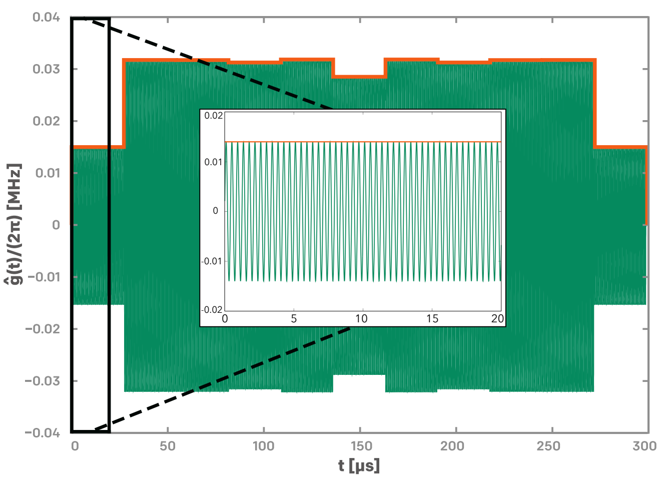
$$\Omega(t) = \Omega_j, \quad \text{for } t_{j-1} \leq t \leq t_j, \quad j = 1, \ldots, N_{\text{seg}}. \quad (73)$$

For this type of pulses our methods are directly applicable with only two minor modifications. (i) For given μ_0, we choose the gate length τ such that $J = \mu_0 \tau/\pi$ is an integer. This way, still requiring that $\Omega(t)$ is an even function with respect to $\tau/2$, we obtain even- or odd-parity pulses, $\hat{g}^{(\pm)}(t) = \Omega(t)\sin(\mu_0 t)$, for J odd or even, respectively. Since J needs to be an integer to obtain the desired symmetry classes, τ can take only discrete values. However, since for quantum computer hardware of practical interest (for instance, Yb-ion quantum computers [25]), the detuning μ_0 is such that J is a large integer (of the order of 1000), the discretization of τ is of no consequence in practice. (ii) The second modification concerns the computation of the matrix M defined in (3) of the main text. For step pulses, we let $M_{pn} \to M_{pj}$, where, including both negative- and positive-parity pulses, we have

$$M_{pj}^{(\pm)} = \int_{t_{j-1}}^{t_j} \sin(\mu_0 t) h_p^{(\pm)}(t)\, dt. \quad (74)$$



Supplementary Figure 2: Amplitude and frequency profiles of pulses used in the implementation of the power optimal, stabilized entangling two-qubit gates. The gate time τ for each of the pulses is $\tau \approx 550.0\mu$s.



Supplementary Figure 3: Fixed-detuning step pulse for $N = 5$, $N_{\rm seg} = 11$, $J = 1434$, $\mu_0/(2\pi) = 2.396\,$MHz, which corresponds to $\tau \approx 299.26\,\mu$s. The thin green line is the step pulse $\hat{g}^{(-)}(t)$; the thick, orange line is the piecewise constant pulse envelope function $\Omega(t)/(2\pi)$.

Defining $\vec{A} = (\Omega_1, \Omega_2, \ldots, \Omega_{N_{\rm seg}})$, all the procedures outlined in Section Supplementary Note 4 can now be applied to construct step pulses.

Supplementary Figure 3 shows an example of a negative-parity step pulse, generated for the same set of motional-mode frequencies and Lamb-Dicke parameters as in the main text Fig. 3**a**. Although Supplementary Fig. 3 shows the negative-parity pulse with the lowest peak-power requirement that we found in the detuning interval from $\mu_0/(2\pi) = 2.2\,$MHz to $\mu_0/(2\pi) = 2.6\,$MHz, we see that this pulse is about 10% higher in peak power than the pulse shown in the main text Fig. 3**a**. This is expected, since fixed-detuning pulses lack the additional degrees of freedom that are associated with being able to modulate the detuning. However, in analogy to the main text Fig. 3**a**, we see that the pulse tends to be relatively flat in amplitude, a feature we observed in all power-optimized pulses we generated in the course of numerous simulations.

There are several reasons why step pulses should be replaced with AMFM pulses. In our opinion the two leading reasons are (i) amplitude-, frequency-, and phase-modulated pulses have lower power requirement as seen when comparing the main text Fig. 3**a** and Supplementary Fig. 3 and (ii) in contrast to the sharp transitions in power levels characteristic for step pulses, amplitude-, frequency-, and phase-modulated pulses have a smooth pulse envelope, which eliminates ringing and the Gibbs phenomenon [26] that accompanies sudden changes in power levels.

In contrast to the straightforward construction of our amplitude-, frequency-, and phase-modulated pulses, finding the optimal pulse for step pulses requires a search in the 4D parameter space consisting of the number of segments, $N_{\rm seg}$, the detuning μ_0, the integer J, and the parity $(\pm)$ of the pulse. While $N_{\rm seg}$ is discrete, and we found that good convergence is already achieved with relatively few segments, in terms of parity there are only two cases to check, and, if τ is pre-specified to a certain value, say, $\tau = 300\,\mu s \pm 1\,\mu s$, the range of J that falls into this interval is not large, and, moreover, J is discrete, searching for the optimal detuning μ_0 requires considerable computational overhead that is avoided using our "single-shot" AMFM approach.

Concluding this subsection, we can say that our new linear algorithm is certainly general enough to encompass the important class of step pulses. Thus, if such pulses are required to run, e.g., existing quantum computers with existing controller hardware which requires step pulses as input, our method can be used to generate these pulses efficiently and directly.

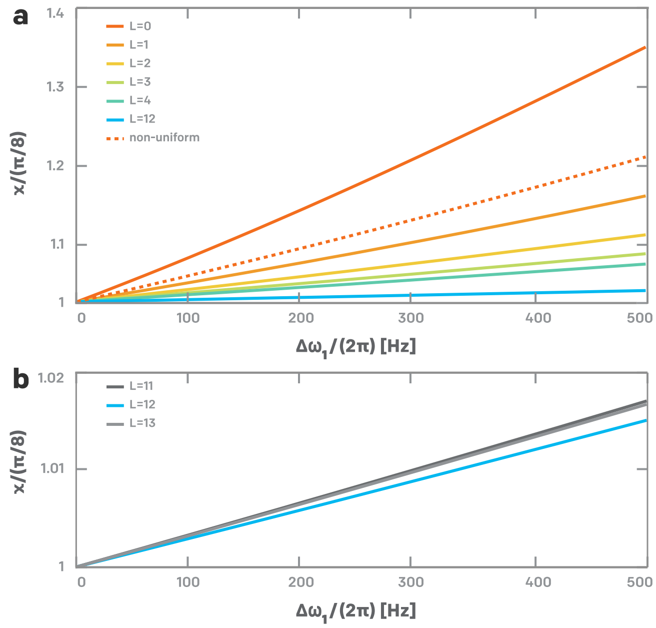
Supplementary Note 13. Efficient arbitrary simultaneously entangling gates

In this section, we show how to use our method in conjunction with the Efficient Arbitrary Simultaneously Entangling (EASE) gate protocol detailed in [27]. To see how this may be achieved, the only thing that is required is to show that the equations to be solved are isomorphic. In particular, the null-space condition (23) is of the same structure as Eq. (2) of [27] and the degree-of-entanglement condition (33) is of the same structure as Eq. (3) of [27], which fully specify the problem of solving for the EASE-gate pulse shapes. The rest of the EASE-gate protocol follows immediately. The resulting pulse shapes can implement up to $N(N-1)/2$ XX gates simultaneously in a short time for a given power budget.

Supplementary Note 14. Sensitivity of the degree of entanglement

In this section we now explore the effects of motional-mode drifts on the gate angle χ. For $N = 5$ ions, two cases are investigated. (i) All modes ω_p drift in unison from 0 Hz to $+2\pi \times 500$ Hz and (ii) individual modes drift independently. For case (ii), we simulated a case in which, chosen randomly, and with random signs of the drift direction, ω_1 drifts from 0 Hz to $+2\pi \times 500$ Hz, ω_2 drifts from 0 Hz to $-2\pi \times 400$ Hz, ω_3 drifts from 0 Hz to $+2\pi \times 300$ Hz, ω_4 drifts from 0 Hz to $-2\pi \times 500$ Hz, and ω_5 drifts from 0 Hz to $+2\pi \times 400$ Hz. Supplementary Figure 4 shows that although all drift amplitudes are substantially smaller than 1 kHz, the effect on the gate angle χ is substantial.

The strong sensitivity of χ with respect to drifts in ω_p is due to the amplification effect of the relatively long pulse duration. In order to compute χ, we have to evaluate the double integral (6). Under the integral we have the term $\sin[\omega_p(t_2 - t_1)]$, and if we replace ω_p by $\omega_p + \Delta\omega_p$, then the $\sin[\omega_p(t_2 - t_1)]$ term becomes, in linear order, $\sin[\omega_p(t_2 - t_1)] + \cos[\omega_p(t_2 - t_1)]\Delta\omega_p(t_2 - t_1)$. Now, while $|\Delta\omega_p|$ is at most $2\pi \times 500$ Hz, which looks small, and indicates that we might be able to neglect the second term, when we multiply the second term with 300 μs, which is the maximum of $t_2 - t_1$, we get $2\pi \times 0.0005\,\text{MHz} \times 300\,\mu\text{s} = 0.94$, which is large. In fact, this term is so large that the linearization approximation breaks down. Therefore, the pulse length is the amplification mechanism and explains the strong sensitivity of χ to relatively small drifts in ω_p. It also underpins the observed sensitivity (see Supplementary Fig. 4) with a detailed qualitative analytical understanding.



Supplementary Figure 4: Projection method: Sensitivity of the gate angle χ to drifts of the motional-mode frequencies ω_p, $p = 1, \ldots, P = N$, for $N = 5$ ions. Shown is the normalized gate angle $|\chi|/(\pi/8)$ as a function of $\Delta\omega_1/(2\pi)$, the drift frequency of motional mode $p = 1$. **a.** Orange solid curve: All five motional modes drift in unison from 0 to $2\pi \times 500$ Hz. Orange dashed curve: The five motional modes drift independently as described in the text. The solid curves document the effect of active χ stabilization against motional-mode drifts as described in the text. **b.** Existence of a "sweet spot" in the number of projected states. For the case chosen ($\tau = 300\,\mu$s and 5 ions) the sweet spot occurs for 12 projected states (blue curves in frames **a** and **b**).

In order to counteract drifts in χ, we suggest to monitor the value of χ continuously and readjust the laser power that drives the XX gate if χ drifts away. This is a valid correction mechanism since the set of equations (1) depends only on the *shape* of the pulse, but not on the pulse *amplitude*. Therefore, without compromising the validity of (1), the power can be continuously adjusted to keep χ within tolerable bounds. Of course, it may be difficult in practice to continuously monitor and readjust χ. Nevertheless, at least in principle, this is a possible correction and stabilization mechanism. In analogy to our moments approach for active stabilization of the α conditions (1), it is also possible to encode active stabilization of χ in the pulse shape itself.

Supplementary Note 15. Single-pulse active stabilization of the degree of entanglement: Projection method

Ideally, to actively stabilize χ_{ij} against ω_p fluctuations, integrated in the pulse-shape construction, we should require

$$\chi_{ij,p}^{(k)} = \frac{\partial^k \chi_{ij}}{\partial \omega_p^k} = 0, \quad k = 1, 2, \ldots, K_\chi, \tag{75}$$

where K_χ is the maximal desired degree of χ stabilization. Since all pulse shapes, regardless of their maximal degree of stabilization K_χ, need to satisfy both the decoupling conditions (1) between the motional modes and the computational states and the degree-of-entanglement condition (7) (where "$\pi/8$" may be replaced by the actual desired degree of entanglement), we may write

$$\chi_{ij,p}^{(k)} = \vec{\Lambda}^T R_p^{(k)} \vec{\Lambda} = 0, \tag{76}$$

where

$$R_{\alpha\beta,p}^{(k)} = (\vec{A}^{(\alpha)})^T S_p^{(k)} \vec{A}^{(\beta)} \tag{77}$$

and

$$S_p^{(k)} = \frac{\partial^k S}{\partial \omega_p^k}. \tag{78}$$

To understand the consequences of (75) [(76), respectively], we spectrally decompose $R_p^{(k)}$ according to

$$R_p^{(k)} = \sum_{\nu=1}^{N_0} \lambda_{\nu,p}^{(k)} |\lambda_{\nu,p}^{(k)}\rangle\langle\lambda_{\nu,p}^{(k)}|, \tag{79}$$

where $\lambda_{\nu,p}^{(k)}$ is the ν-th eigenvalue of $R_p^{(k)}$ and $|\lambda_{\nu,p}^{(k)}\rangle$ is the corresponding eigenvector. Expanding $\vec{\Lambda}$ into the eigenstates of $R_p^{(k)}$, i.e.,

$$\vec{\Lambda} = \sum_{\nu=1}^{N_0} c_{\nu,p}^{(k)} |\lambda_{\nu,p}^{(k)}\rangle, \tag{80}$$

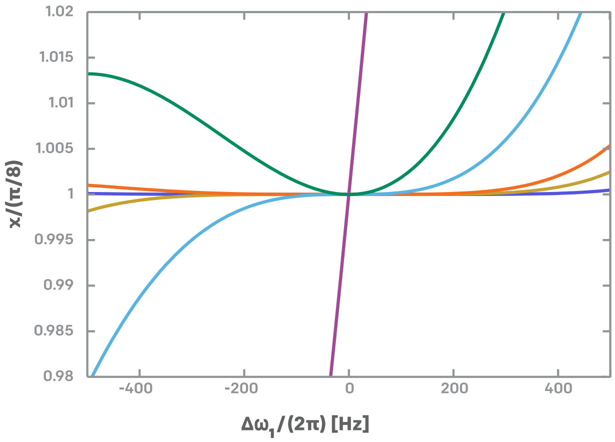
the stabilization condition (76) may then be written as

$$\chi_{ij,p}^{(k)} = \sum_{\nu=1}^{N_0} |c_{\nu,p}^{(k)}|^2 \lambda_{\nu,p}^{(k)} = 0. \tag{81}$$

Equation (81) brings out the problem: The condition (81) can be satisfied only if not all of the eigenvalues $\lambda_{\nu,p}^{(k)}$ have the same sign. However, we can prove analytically (the proof is lengthy and not shown here) and confirmed numerically, that, for instance, $R_p^{(k=1)}$ is a definite matrix for all p, i.e., the eigenvalues of $R_p^{(k=1)}$ are all non-zero and have the same sign, which makes it impossible to satisfy (81) for $k = 1$. We did, however, notice that only a few of the eigenvalues of $R_p^{(k=1)}$ are particularly large in absolute magnitude, which may be the ultimate reason for the strong sensitivity of χ_{ij} in linear order (see Supplementary Fig. 4). This observation suggests a strategy for actively stabilizing χ_{ij} against ω_p-fluctuations: Projecting those components of the spectra of $R_p^{(k)}$ out of the null-space of M (which can be assumed to already include stabilization of (1) against ω_p-fluctuations) that correspond to the eigenvalues with the largest absolute values. If we project out L such components from each of the $R_p^{(k)}$ matrices, this leaves us with a null space of $N_0' = N_0 - PK_\chi L$ dimensions that now, to a large degree, actively stabilizes χ_{ij} against ω_p-fluctuations. Following this projection step, we now use the techniques presented in Supplementary Note 4, applied to the reduced null space of N_0' dimensions, to satisfy the degree-of-entanglement condition with the smallest possible average power.

To illustrate this technique, and focusing on the case of uniform drift of the motional modes from 0 to 500 Hz as defined above, we present in Supplementary Fig. 4**a** the result of projecting 1,2,3,4, and 12 states from the null space that correspond to the eigenvalues with largest absolute values of $R_p^{(1)}$, $p = 1, \ldots, 5$. We see that already for a single projected state we achieve noticeable stabilization that improves further for 2, 3, and 4 projected states. This improvement continues if more states are projected, reaching an optimum ("sweet spot") for 12 projected states. This is illustrated in Supplementary Fig. 4**b**, which shows the normalized χ for 11, 12, and 13 projected states. Therefore, while we found that projecting relatively few states always results in improved active stabilization, "over-projection" should be avoided, since it is both costly in power and does not improve χ stabilization any further. In fact, as expected, active χ stabilization, in analogy with stabilizing α, requires increased levels of power. For example, for the case shown in Supplementary Fig. 4, the projection of 1, 2, 3, 4, and 12 states requires power levels of 1.5, 1.7, 2.0, 2.1, and 3.3 with respect to the power level without projection. But we also see that projection is relatively inexpensive compared with the significant amount of stabilization gained.

The projection technique works for all orders $k \geq 1$ of $R_p^{(k)}$. However, not much is gained by continuing the projection beyond $k = 1$. The reason is the following. In our example, the best result, obtained by projecting the first 12 states, brings down the variation in the relative χ from 35% to just 1.5%. These 1.5%, however, are mostly due to the residual slope of the first-order stabilization, so that second-order stabilization would not contribute much, other than computational effort and power expended. We see this in the following way. The slope of the first-order stabilization for 12 projected states at 0 motional-mode drift is 3×10^{-5} /Hz. Therefore, at 500 Hz motional-mode drift, the variation in the relative value of χ is 0.015, i.e. 1.5%. This is exactly the amount we read off in Supplementary Fig. 4**b** for the case of 12 projected



Supplementary Figure 5: Two-pulse moments approach: Sensitivity of the gate angle χ to uniform drifts of the motional-mode frequencies ω_p, $p = 1, \ldots, P = N$ for $N = 5$ ions as a function of the drift frequency $\Delta\omega_1$ of the first motional mode for the first six stabilization orders K_χ. Different color lines correspond to different stabilization orders K_χ. Purple: $K_\chi = 0$; green: $K_\chi = 1$; cyan: $K_\chi = 2$; orange: $K_\chi = 3$; yellow: $K_\chi = 4$; blue: $K_\chi = 5$. Compared with Supplementary Fig. 4, the two-pulse moments method provides significantly improved χ stabilization.

states. Therefore, the residual variation is mostly due to the first order, and stabilizing the second order will have a negligible effect. Nevertheless, as shown in Supplementary Fig. 4**a**, and given an unstabilized variation of 35%, active first-order stabilization via projection, which brings this variation down to about 1.5% (see Supplementary Fig. 4**b**), is already significant for the stabilization of quantum-computer operation. While, as discussed above, it is not possible to implement the moments strategy (75) with a single pulse, working with two different pulses, each directed at a different one of the two ions participating in the gate, (75) can in fact be realized and results in the moments methods described in the following two sections.

Supplementary Note 16. Two-pulse active stabilization of the degree of entanglement: Moments method

While it is impossible to satisfy (75) using identical pulses directed at both ions i and j, the condition (75) can be satisfied using different pulses directed at ions i and j. Let $g^{(i)}(t)$ and $g^{(j)}(t)$ be the pulse functions directed at ions i and j, respectively, and let $\hat{\vec{F}}$ and $\hat{\vec{G}}$ be the null-space expansion amplitudes of $g^{(i)}(t)$ and $g^{(j)}(t)$, respectively. Then, condition (5) in the main text implies

$$\chi = \frac{\pi}{8} = \hat{\vec{F}}^T R \hat{\vec{G}}, \tag{82}$$

which has to be solved under the conditions [see (75)]

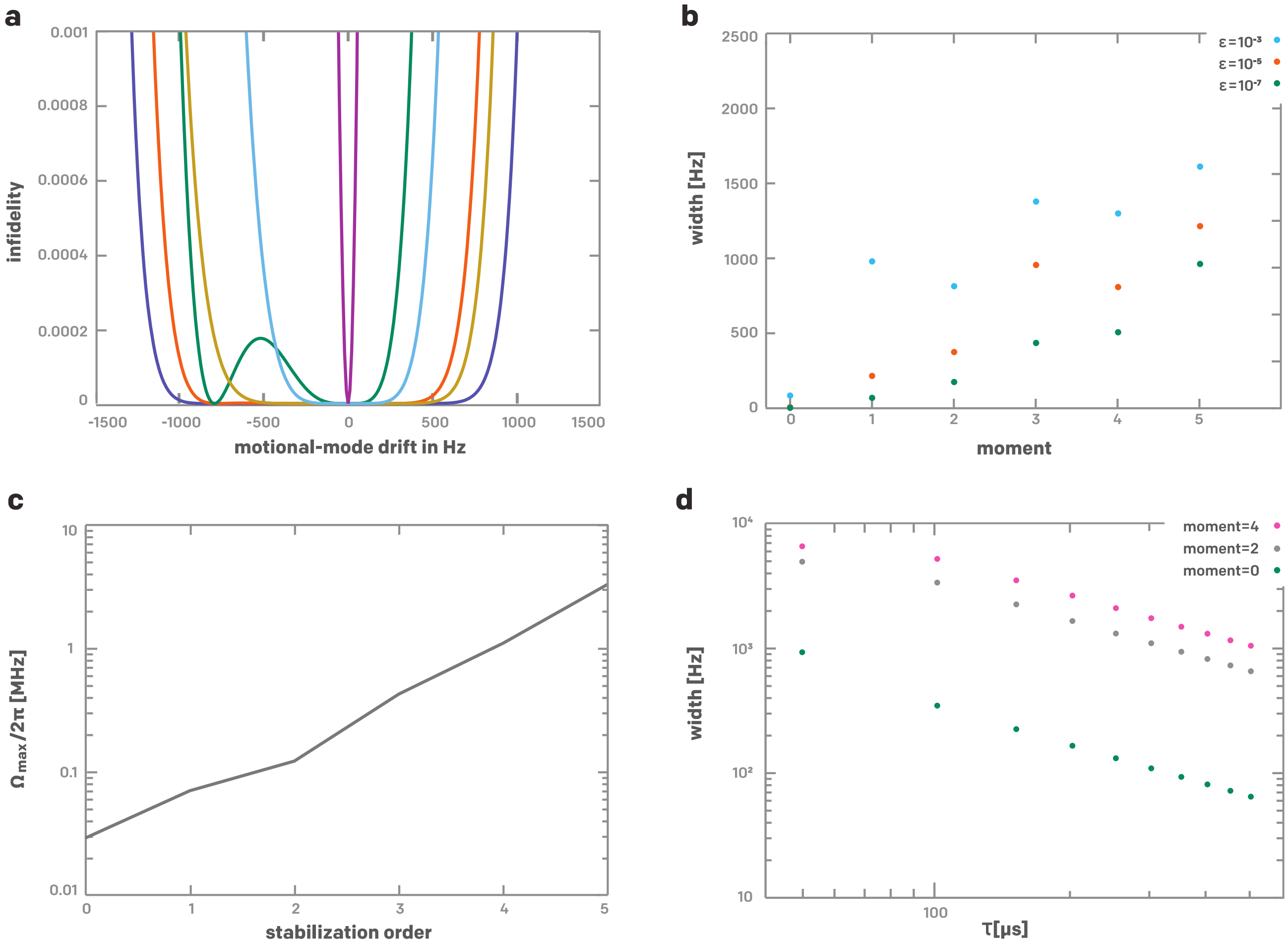
$$\vec{F}^T R_p^{(k)} \vec{G} = 0, \quad p = 1, \ldots, P, \quad k = 1, \ldots, K_\chi, \tag{83}$$

where $\vec{F}$ and $\vec{G}$ are any un-normalized versions of $\hat{\vec{F}}$ and $\hat{\vec{G}}$. Unlike in the single-pulse case discussed in Supplementary Note 15, where, due to the observed definiteness of $R_p^{(1)}$, it was impossible to satisfy (75), working with two different pulses, $\vec{F}$ and $\vec{G}$, (83) can be satisfied as soon as $\vec{F}$ and $\vec{G}$ are orthogonal to each other with respect to $R_p^{(k)}$.

An explicit solution of (82) and (83) can be constructed in the following way. We start by choosing $\vec{G}$ to be the power-optimal pulse as computed in the single-pulse case described in Supplementary Note 4. Then, we construct the space $\mathcal{P}$, which is spanned by the vectors $\vec{v}_p^{(k)} = R_p^{(k)}\vec{G}$, $p = 1, \ldots, P$, $k = 1, \ldots, K_\chi$. We also define the space $\mathcal{Q}$, which is the orthogonal complement of $\mathcal{P}$ with respect to the null space. With these definitions, taking $\vec{F}$ out of $\mathcal{Q}$, we obtain the most power-optimal solution for $\vec{F}$ by choosing $\vec{F} = \hat{Q}\vec{G}$, where $\hat{Q}$ projects into the $\mathcal{Q}$ space. At this point the normalizations of $\vec{F}$ and $\vec{G}$ are still two free parameters. We use the first scaling freedom to obtain $|\hat{\vec{F}}| = |\hat{\vec{G}}|$ (symmetric pulse power) and use the second scaling freedom to satisfy (82).

For the same case of uni-directional ω_p drift as used in Supplementary Fig. 4, the performance of active two-pulse χ stabilization is illustrated in Supplementary Fig. 5 and can be compared with the performance of the single-pulse projection method illustrated in Supplementary Fig. 4. We see that while the projection method at best cuts the relative χ drift error down to 1.5% at 500 Hz drift, the relative error in the case of $K_\chi = 5$ stabilization orders is smaller than 1 permille even at 500 Hz drift. In addition, while in the projection method the stabilization is of linear order around zero motional-mode drift, the moments method produces favorable non-linear stabilization of order $K_\chi + 1$ around zero motional-mode drift (see Supplementary Fig. 5). The power of the moments method is best appreciated in Supplementary Fig. 5 by comparing the behavior of the unstabilized χ (the nearly vertical, purple line in Supplementary Fig. 5 around zero motional-mode drift; $K_\chi = 0$) with the shape of χ as a function of motional-mode drift for different $K_\chi \geq 1$. Even for small $K_\chi \geq 1$, substantial stabilization is observed.

In Supplementary Fig. 6 we show a summary of various aspects of the two-pulse active χ stabilization method. Supplementary Figure 6**a** shows the infidelity of the two-pulse active-stabilization method as a function of mode-frequency drift. We see that infidelities $\ll 10^{-3}$ can be achieved for motional-mode frequency drifts in the range ± 200 Hz for all $K_\chi \geq 1$ and and an infidelity of $\ll 10^{-4}$ can be achieved over a motional-mode frequency drift in the range ± 500 Hz for $K_\chi = 5$. Supplementary Figure 6**b** shows a summary of frequency widths of χ stabilization for infidelity cut-offs of 10^{-3}, 10^{-5}, and 10^{-7},



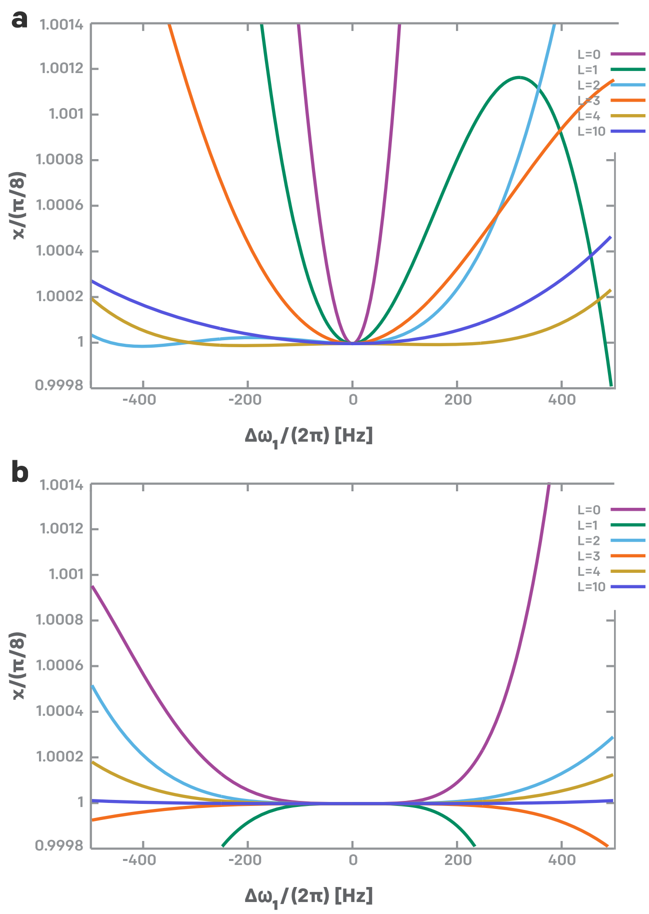
Supplementary Figure 6: Stabilization of the control pulses. **a**. Infidelity (see SI section Supplementary Note 8 for detail) as a function of the motional-mode frequency drift Δf. All mode frequencies were drifted according to $\omega_p \mapsto \omega_p + 2\pi\Delta f$. **b**. The width of the infidelity curves in **a** for various error tolerances $\epsilon = 10^{-3}, 10^{-5}$, and 10^{-7}, as a function of the highest moment K_χ of stabilization. **c**. The maximal power requirement $\max_t |g_K(t)|$ of the control pulses as a function of the highest moment K_χ of stabilization. The power requirement suggests an exponential scaling of power in the order of stabilization K_χ. **d**. Width of the infidelity curves for various different orders of stabilization $K_\chi = 0, 2,$ and 4, as a function of the gate duration τ for a fixed error tolerance level $\epsilon = 10^{-3}$. The data suggests $\sim 1/\tau$ scaling of the width.

respectively. On average we observe a linear increase of frequency width with the χ stabilization order. While these results are promising, they do come with a price. Supplementary Figure 6**c** shows the power requirement as a function of stabilization order K_χ. In constrast with the linear power scaling of α stabilization shown in main text Fig. 4**c**, Supplementary Fig. 6**c** indicates that the required pulse power in the case of χ stabilization increases exponentially with the stabilization order. This, however, should not discourage us from using the moments method for active two-pulse stabilization of χ, since, as shown in Supplementary Figs. 5 and 6, significant χ stabilization is already achieved for relatively small K_χ. Supplementary Figure 6**d** shows the scaling of frequency width for several stabilization orders K_χ as a function of gate duration τ. Similar to the results shown in main text Fig. 4**d**, we observe that the frequency width is inversely proportional to the gate duration.

Supplementary Note 17. Two-pulse active stabilization of the degree of entanglement: Hybrid method

Combining the most advantageous features of both the moments and projection methods, we arrive at the hybrid method. In this method, for desired stabilization order K_χ, we first construct the pulse $\hat{\vec{G}}$ as described in section Supplementary Note 15, i.e., as a single pulse in a null space of dimension $N_0' = N_0 - PL$, where we projected out those L eigenvectors from each matrix $R_p^{(1)}$, $p = 1, \ldots, P$, that correspond to the L eigenvalues with the largest absolute values. We then proceed with the construction of the pulse $\hat{\vec{F}}$ as described in section Supplementary Note 16. Thus, even before construction of the pulse $\hat{\vec{F}}$, we eliminate those null-space components from the pulse $\hat{\vec{G}}$ that potentially produce the most sensitivity in χ. As documented in Supplementary Fig. 7, and compared with the single-pulse projection method (see Supplementary Note 15) and the bare two-pulse moments method (see Supplementary Note 16), the hybrid method yields the best results in terms of active χ stabilization.

Supplementary Figure 7**a**, for the same case of uniform mode-frequency drift as used in Supplementary Figs. 4, 5, and 6, shows the result of χ stabilization for $K_\chi = 1$ and L vectors with the largest absolute values of their eigenvalues projected out. We see that already for the case $K_\chi = 1$, compared with the bare moment method for $K_\chi = 1$ (see Supplementary Fig. 5), the gain in stabilization is substantial. This observation is important since, because of the exponential power cost of the two-pulse method in K_χ, it is advantageous, for a given stabilization target, to stabilize with the smallest possible K_χ. Supplementary Figure 7**b** shows that with a slightly larger K_χ [$K_\chi = 3$ in Supplementary Fig. 7**b**], a substantial broadening of the stabilization region can be achieved. Supplementary Figure 7**b** also shows that by projecting out $L = 10$ vectors, χ stabilization on the level of 10^{-5} can be achieved for 500 Hz motional-mode drift.



Supplementary Figure 7: Performance of the hybrid method. **a**. For $K_\chi = 1$, $L = 0, 1, 2, 3, 4$, and $L = 10$ vectors with the largest absolute value of their corresponding eigenvalues are projected out. **b**. Same as **a**. but with $K_\chi = 3$. **b**. shows that compared with **a**. larger K_χ results in a broader stabilization region, which is widened even further by projection.

Supplementary Note 18. Broadband sequence

A host of compensation pulse sequences that mitigate the errors in the single-qubit gate are known (see [28] and the references therein) and similar techniques can indeed be used to mitigate the errors in χ_{ij} that arise from, e.g., relative offsets in η_i^p or $g(t)$. We refer interested readers to [29, 30], where a variety of broadband behaviors have been explored. Below, we show an example for completeness.

Consider a broadband behavior of the Solovay-Kitaev (SK) sequence in [28] to compensate for the inexact χ_{ij} up to first order. Higher orders or other compensation techniques, such as those that rely on Suzuki-Trotter sequences, may straightforwardly be employed. Typically, the SK compensation sequence is discussed in the context of single-qubit operators, where, for small error strength ϵ,

$$R(\theta, 0) - R(2\pi(1+\epsilon), -\phi_{\rm SK})R(2\pi(1+\epsilon), \phi_{\rm SK})R(\theta(1+\epsilon), 0) = O(\epsilon^2), \tag{84}$$

where

$$R(\theta, \phi) = \exp\{-i\theta[\cos(\phi)\sigma_x + \sin(\phi)\sigma_y]/2\} \tag{85}$$

and

$$\phi_{\rm SK} = \cos^{-1}(-\theta/4\pi). \tag{86}$$

A straightforward extension to the two-qubit $\mathrm{XX}(\theta)$ gate may be done to result in

$$\mathrm{XX}(\theta) - \mathrm{X}\bar{\phi}(2\pi(1+\epsilon))\mathrm{X}\phi(2\pi(1+\epsilon))\mathrm{XX}(\theta(1+\epsilon)) = O(\epsilon^2), \tag{87}$$

where

$$\mathrm{X}\phi(\theta) = (1 \otimes \mathrm{R}_z(\phi_{\rm SK}))\mathrm{XX}(\theta)(1 \otimes \mathrm{R}_z(-\phi_{\rm SK})),$$
$$\mathrm{X}\bar{\phi}(\theta) = (1 \otimes \mathrm{R}_z(-\phi_{\rm SK}))\mathrm{XX}(\theta)(1 \otimes \mathrm{R}_z(\phi_{\rm SK})), \tag{88}$$

and

$$\mathrm{R}_z(\phi) = \exp(-i\theta\sigma_z/2). \tag{89}$$

The choice of application of R_z gates on the second qubit is arbitrary, and can indeed instead be performed on the first qubit without loss of generality. Because the errors in χ_{ij} incur in one well-defined direction of $\sigma_x\sigma_x$ in the 15-dimensional hyper-Bloch sphere, the single-qubit compensation-pulse techniques become straightforwardly applicable.

Supplementary Note 19. Direct implementation of Fourier-basis pulse function

According to [17], in the Lamb-Dicke regime, the interaction Hamiltonian for the ion-chain system, subjected to a dual-tone, symmetric blue- and red-sideband beam with detuning $\pm\mu$, in the x basis is

$$H_{\text{dual-tone}}(t) = \sum_{i=1}^{N}\sum_{p=1}^{P}\Omega_i(t)\eta_p^i \sin(\mu t)(a_p e^{-i\omega_p t} + a_p^\dagger e^{i\omega_p t})\sigma_x^i, \tag{90}$$

where a_p and $a_p^\dagger$ are the annihilation and creation operators of the pth motional mode, respectively. Consider now a multi-tone beam with amplitudes $\Omega_{i,n}(t)$ and detuning frequencies $\pm\mu_n$, where $n = 1, 2, .., N_A$. This results in the Hamiltonian

$$H_{\text{multi-tone}}(t) = \sum_{i=1}^{N}\sum_{p=1}^{P}\sum_{n=1}^{N_A}\Omega_{i,n}(t)\eta_p^i \sin(\mu_n t) (a_p e^{-i\omega_p t} + a_p^\dagger e^{i\omega_p t})\sigma_x^i. \tag{91}$$

Define

$$f_{ip}(t) = \eta_p^i g_i(t), \tag{92}$$

where

$$g_i(t) = \sum_{n=1}^{N_A}\Omega_{i,n}(t)\sin(\mu_n t), \tag{93}$$

where N_A is chosen sufficiently large to achieve convergence. Inserting (92) in (91), we obtain

$$H_{\text{multi-tone}}(t) = \sum_{i=1}^{N}\sum_{p=1}^{P} f_{ip}(t)(a_p e^{-i\omega_p t} + a_p^\dagger e^{i\omega_p t})\sigma_x^i, \tag{94}$$

which induces the system evolution over the gate time τ described by

$$U_{\text{multi-tone}}(t) = \exp\left\{ -i\int_0^\tau dt H_{\text{multi-tone}}(t) - \frac{1}{2}\int_0^\tau dt_2 \int_0^{t_2} dt_1 [H_{\text{multi-tone}}(t_2), H_{\text{multi-tone}}(t_1)]\right\}, \tag{95}$$

as shown in [17] using Magnus' formula. Inserting (94) in (95), together with (92) and (93), we obtain, up to a global phase,

$$U_{\text{multi-tone}}(t) = \exp\left\{ -i\left[\sum_{i=1}^{N}\sum_{p=1}^{P}\left(\alpha_{ip}\eta_p^i a_p + \alpha_{ip}^*\eta_p^i a_p^\dagger\right)\sigma_x^i\right] + i\sum_{i,j=1;i\neq j}^{N}\chi_{ij}\sigma_x^i\sigma_x^j\right\}, \tag{96}$$

where

$$\alpha_{ip} = \int_0^\tau g_i(t)e^{-i\omega_p t}dt, \tag{97}$$

α_{ip}^* denotes its complex conjugate, and

$$\chi_{ij} = \sum_{p=1}^{P}\eta_p^i\eta_p^j\int_0^\tau dt_2\int_0^{t_2}dt_1 g_i(t_2)g_j(t_1)\sin[\omega_p(t_2-t_1)]. \tag{98}$$

Comparing (97) and (98) with main text Eqs. (3) and (4), respectively, together with (93), and assuming $g_i(t) = g_j(t) = g(t) = \sum_{n=1}^{N_A} A_n \sin(2\pi n t/\tau)$, as we did in the main text, we see that $\Omega_{i,n}(t) = A_n$ and $\mu_n = 2\pi n/\tau$ implements the pulse function that implements the desired XX gate.

[1] Y. Nam *et al.*, Ground-state energy estimation of the water molecule on a trapped ion quantum computer. *npj Quantum Inf.* **6**, 33 (2020).

[2] Y. Nam and D. Maslov, Low cost quantum circuits for classically intractable instances of the Hamiltonian dynamics simulation problem. *npj Quant. Inf.* **5**, 44 (2019).

[3] G. E. Crooks, Performance of the quantum approximate optimization algorithm on the maximum cut problem. https://arxiv.org/abs/1811.08419 (2018).

[4] Y. Nam, Y. Su, D. Maslov, Approximate quantum Fourier transform with $O(n\log(n))$ T gates. *npj Quantum Inf.* **6**, 26 (2020).

[5] T. G. Draper, S. A. Kutin, E. M. Rains, K. M. Svore, A logarithmic-depth quantum carry-lookahead adder. *Quant. Inf. Comp.* **6**, 351–369 (2006).

[6] M. Benedetti, D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, A. Perdomo-Ortiz, A generative modeling approach for benchmarking and training shallow

quantum circuits. *npj Quant. Inf.* **5**, 45 (2019).

[7] R. Babbush *et al.*, Encoding electronic spectra in quantum circuits with linear T complexity. *Phys. Rev. X* **8**, 041015 (2018).

[8] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, M. Troyer, Elucidating reaction mechanisms on quantum computers. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7555–7560 (2017).

[9] M. X. Goemans, D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM* **42**, 1115-1145 (1995).

[10] A. M. Childs, D. Maslov, Y. Nam, N. J Ross, Y. Su, Toward the first quantum simulation with quantum speedup. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9456–9461 (2018).

[11] V. V. Shende, I. L. Markov, S. S. Bullock, Minimal universal two-qubit controlled-NOT-based circuits. *Phys. Rev. A* **69**, 062321 (2004).

[12] J. Lee, D. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, R. Babbush, Even more efficient quantum computations of chemistry through tensor hypercontraction. https://arxiv.org/abs/2011.03494 (2020).

[13] L. K. Grover, Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79**, 325 (1997).

[14] F. A. Aloul, A. Ramani, I. L. Markov, K. A. Sakallah, Solving difficult SAT instances in the presence of symmetry. *Proc. Des. Automat. Conf.* 731–736 (2002).

[15] D. Maslov, On the advantages of using relative phase Toffolis with an application to multiple control Toffoli optimization. *Phys. Rev. A* **93**, 022311 (2016).

[16] K. Mølmer, A. Sørensen, Multiparticle Entanglement of Hot Trapped Ions, *Phys. Rev. Lett.* **82**, 1835-1838 (1999).

[17] S.-L. Zhu, C. Monroe, L.-M. Duan, Arbitrary-speed quantum gates within large ion crystals through minimum control of laser beams. *Europhys. Lett.* **73**, 485 (2006).

[18] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, D. M. Meekhof, Experimental Issues in Coherent Quantum-State Manipulation of Trapped Atomic Ions. *J. Res. Natl. Inst. Stand. Technol.* **103**, 259–328 (1998).

[19] T. Choi, S. Debnath, T. A. Manning, C. Figgatt, Z.-X. Gong, L.-M. Duan, and C. Monroe, Optimal Quantum Control of Multimode Couplings between Trapped Ion Qubits for Scalable Entanglement. *Phys. Rev. Lett.* **112**, 190502 (2014).

[20] P. H. Leung, K. A. Landsman, C. Figgatt, N. M. Linke, C. Monroe, K. R. Brown, Robust 2-qubit gates in a linear ion crystal using a frequency-modulated driving force. *Phys. Rev. Lett.* **120**, 020501 (2018).

[21] T. J. Green, M. J. Biercuk, Phase-modulated decoupling and error suppression in qubit-oscillator systems. *Phys. Rev. Lett.* **114**, 120502 (2015).

[22] K. Wright *et al.*, Benchmarking an 11-qubit quantum computer. *Nat. Commun.* **10**, 5464 (2019).

[23] C. Marquet, F. Schmidt-Kaler, D. F. V. James, Phonon-phonon interactions due to non-linear effects in a linear ion trap. *Appl. Phys. B* **76**, 199–208 (2003).

[24] Y. Wu, S.-T. Wang, L.-M. Duan, Noise Analysis for High-Fidelity Quantum Entangling Gates in an Anharmonic Linear Paul Trap. *Phys. Rev. A* **97**, 062325 (2018).

[25] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, C. Monroe, Demonstration of a small programmable quantum computer with atomic qubits. *Nature* **536**, 63–66 (2016).

[26] E. Hewitt, R. E. Hewitt, The Gibbs-Wilbraham phenomenon: An episode in Fourier analysis, *Archive for History of Exact Sciences* **21**, 129–160 (1979).

[27] N. Grzesiak, R. Blümel, K. Beck, K. Wright, V. Chaplin, J. Amini, N. Pisenti, S. Debnath, J.-S. Chen, Y. Nam, Efficient Arbitrary Simultaneously Entangling Gates on a trapped-ion quantum computer. https://arxiv.org/abs/1905.09294 (2019); *Nat. Commun.* **11:2963** (2020), https://doi.org/10.1038/s41467-020-16790-9

[28] J. T. Merrill, K. R. Brown, Progress in compensating pulse sequences for quantum computation. https://arxiv.org/abs/1203.6392 (2012).

[29] Svetoslav S. Ivanov, Nikolay V. Vitanov, Composite two-qubit gates. *Phys. Rev. A* **92**, 022333 (2015).

[30] Daniel C. Murphy, Kenneth R. Brown, Controlling error orientation to improve quantum algorithm success rates. *Phys. Rev. A* **99**, 032318 (2019).