Quantum-enhanced bosonic learning machine

Chi-Huan Nguyen,1 Ko-Wei Tseng,1 Gleb Maslennikov1, H. C. J. Gan,1 and Dzmitry Matsukevich1,2

1Centre for Quantum Technologies, National University of Singapore, 3 Science Dr 2, 117543, Singapore
2Department of Physics, National University of Singapore, 2 Science Dr 3, 117551, Singapore
(Dated: April 12, 2021)

Quantum processors enable computational speedups for machine learning through parallel manipulation of high-dimensional vectors [1]. Early demonstrations of quantum machine learning have focused on processing information with qubits [2–8]. In such systems, a larger computational space is provided by the collective space of multiple physical qubits. Alternatively, we can encode and process information in the infinite dimensional Hilbert space of bosonic systems such as quantum harmonic oscillators [9–11]. This approach offers a hardware-efficient solution with potential quantum speedups to practical machine learning problems. Here we demonstrate a quantum-enhanced bosonic learning machine operating on quantum data with a system of trapped ions. Core elements of the learning processor are the universal feature-embedding circuit that encodes data into the motional states of ions, and the constant-depth circuit that estimates overlap between two quantum states. We implement the unsupervised K-means algorithm to recognize a pattern in a set of high-dimensional quantum states and use the discovered knowledge to classify unknown quantum states with the supervised k-NN algorithm. These results provide building blocks for exploring machine learning with bosonic processors.

Machine learning techniques excel at finding patterns in data that are usually encoded as high dimensional feature vectors. As the size of these vectors grows, processing them classically requires enormous computational resources [12]. Recently proposed quantum algorithms show evidences of computational speedups, that comes from the exploitation of large Hilbert space of quantum systems [12–15]. Demonstrations of quantum machine learning algorithms implemented on discrete variables platforms have shown tremendous potential [2–7]. However, the number of gates and qubits must be increased for solving practical problems on near-term quantum processors.

An alternative approach is to utilize the infinite-dimensional Hilbert space of bosonic systems as the feature space for quantum machine learning [16, 17]. One advantage of the bosonic architecture is that a single operation, such as squeezing or displacement, is formally equivalent to processing infinitely-dimensional feature vectors [17]. Furthermore, it allows for efficient use of physical resources. For discrete-variable schemes, approximately log(n) physical qubits are required to encode an n-bit data string, whereas in the alternate approach, one bosonic mode is sufficient. As an example, in the ideal case the motion of a single trapped ion suffices to encode 3 infinitely dimensional feature vectors, whereas the internal state of an ion typically only encodes a single qubit [10].

In this letter, we report a bosonic learning machine that consists of two building blocks for encoding and processing either quantum or classical data (see Fig. 1). First, a universal embedding circuit implements a feature map \( \phi \) of classical data \( x = (x_1, x_2, ..., x_n) \) to a quantum state \( |\phi(x)\rangle \). The map chosen in this work is amplitude encoding \( |\phi_{\text{amp}}(x)\rangle = 1/||x||_2 \sum_{j=1}^{||x||_2} |j - 1\rangle \), where \( \{|j\}\rangle \) are the Fock states, \( n \) denotes the dimension of the data, and \( ||x||_2 = \sqrt{\sum_{j=1}^{n} x_j^2} \) is the Euclidean norm. The universal embedding circuit \( \phi \) can be associated with a state preparation gate sequence \( U_\phi(x) \), that prepares an arbitrary state \( |\phi(x)\rangle \) satisfying \( U_\phi(x)|0\rangle = |\phi(x)\rangle \) [17, 18].

Machine learning algorithms frequently rely on distance measures that can quantify the similarity between feature vectors [19]. Here, we employ the Hilbert-Schmidt distance defined as \( D_{\text{HS}}(|\phi(x)\rangle, |\phi(x')\rangle) = \sqrt{2 - 2(|\langle\phi(x)|\phi(x')\rangle|^2)} \), where \( \{|\langle\phi(x)|\phi(x')\rangle|^2 \} \) is the overlap between two quantum states \( |\phi(x)\rangle \) and \( |\phi(x')\rangle \) [3]. The overlap measurement, reported throughout this work, can be efficiently realized with a constant-depth circuit implemented by a SWAP test [20, 23].

In this work, the bosonic learning machine comprises two trapped \( ^{171}\text{Yb}^+ \) ions in a linear radio-frequency Paul trap, employing three motional modes denoted \( A, B, \) and \( C \). We start the experiment by preparing all the modes in the ground state of motion. Two of the modes, \( B \) and \( C \), are used for storing information while \( A \) acts as an auxiliary mode required for implementing the SWAP test. The hyperfine states \( |g\rangle, |e\rangle \) of one ion mediate interaction between the motional modes, while the other ion is prepared in the \( ^2F_{7/2} \) metastable state and does not interact with optical fields during the experiment (See Fig. 1). The Jaynes-Cumming interaction between the internal state and the motional modes enables deterministic preparation of an arbitrary state with dimension \( n \) in \( n - 1 \) steps, using methods presented in [21, 25]. Alternatively, we use displacement and squeezing operations to prepare coher-
between two states is proportional to the probability $P$ of a projection measurement of the internal state. The overlap $U$ of gates $U$ can be implemented with controlled beam splitter gates.

Two Raman beams (R1, R2) couple the motion in the radial directions with the internal states prepared in mode $B$ (creation) operator of the motional modes, $g_{20, 21, 23}$. Here $T$ is the Hamiltonian of the system.

The internal states $|g\rangle, |e\rangle$ are expressed in the Fock basis $\{|n\rangle\}_{n\geq0}$, $\{|m\rangle\}_{m\geq0}$, and $\{|x\rangle\}$ are frequencies of the motional modes $A$, $B$, and $C$ which denote the out-of-phase motion in the $y$ direction, out-of-phase and in-phase motion in the $x$ direction respectively. The arrows represent the Raman beams.

To detect the internal state in $|g\rangle$: $|\langle\phi(x)|\phi(x')\rangle|_C^2 = |1 - 2F_g|$, Fig. 1]. The SWAP test duration is independent from the dimension of the data and is significantly shorter than the coherence times of the motional modes $|23\rangle$.

To highlight applications of our bosonic learning machine, we demonstrate the unsupervised $K$-means clustering and supervised $k$-NN algorithms using quantum states as the input data. We note that the input data may also be classical as it can be mapped to quantum states by using the universal embedding circuit. We begin with the description of implementing the unsupervised $K$-means clustering algorithm. Consider two parties Alice and Bob. Alice prepares and sends Bob identical copies of $M$ quantum states. The quantum states can be expressed in the Fock basis $|d_m\rangle = \sum_{j=0}^{n-1} x_{m,j+1} |j\rangle$, where $m$ runs from 1 to $M$. In this scenario, Bob is
We choose the parameters as $r$ and $|t|$. The quantum advantage comes from the reduction of complexity of centroids corresponding to the $k$-th cluster. The classification results of $K$-means algorithm into three clusters are labeled by different colors: cluster 1 (red), cluster 2 (blue), and cluster 3 (purple). The algorithm assigns $H_{sqz}$ to cluster 1, $H_{Fock}$ to cluster 2, and $H_{coh}$ to cluster 3. Inset shows the process of updating the centroids visualized in the $x_1$-$x_2$ scatter plot. The error bar denotes the one $\sigma$ fit error of the state characterization (see Methods).

FIG. 2: $K$-means clustering results. a. Measurement of overlap $|\langle d_m | c_{lk} \rangle|^2$ between the data states and the $k$-th centroid $|c_{lk}\rangle$ of the assignment step $l$. The data states are the squeezed vacuum states ($H_{sqz}$), superposition of $|0\rangle$ and $|1\rangle$ Fock states ($H_{Fock}$), and coherent states ($H_{coh}$). Top row shows the step-by-step assignment of states into clusters, visualized in the $x_1$-$x_4$ scatter plot. b. The scatter plot matrix of the centroids (crosses) and data states: $H_{sqz}$ (circles), $H_{Fock}$ (squares), and $H_{coh}$ (diamonds). The assignment results after each step the centroids, truncated to dimension 5, are updated and approximately synthesized using the embedding circuit. We characterize the state generation through measurements of Rabi oscillations on the $|g\rangle|n+1\rangle$ to $|e\rangle|n\rangle$ red-sideband transition (see Methods).

Figure 2a shows the overlap measurements between the data states and the centroids. We used on average 700 experimental runs for each measurement. With this particular choice of the initial centroids, the algorithm attains a stationary assignment after four steps. The final classification result is given in Fig. 2b, where the scatter plot of the amplitudes $x_j$ of the data states are shown for $j$ up to 5. The algorithm associates cluster 1 to $H_{sqz}$, cluster 2 to $H_{Fock}$, and cluster 3 to $H_{coh}$. The SWAP test measurements, albeit imperfect, produce classification results that agree perfectly with the prior knowledge about the states sent to Bob. Stationary assignment of the centroids is evident from the step-by-step evolution shown in the inset of Fig. 2b. The assignment results after each step can be found in Extended Data.
unknown trial states, by using the three labelled clusters from the K-means results as training states. The traditional approach to solving the problem of classifying unknown states is to perform a complete tomography, which requires a number of measurements that grows exponentially with the state dimension $n$ \cite{26}. However, if complete information regarding the state is unnecessary, $k$-NN algorithm could offer an efficient alternative, especially when $n$ is large. With a constant depth SWAP test, computational complexity of the $k$-NN becomes independent of $n$.

The $k$-NN algorithm requires estimation of the overlap between each trial state and all training states. We identify the $k$ states having highest overlap with the trial state. The trial state is then assigned to the cluster that labels the majority of $k$ states. Figure 3 shows the experimental results of $k$-NN classification with $k = 7$ for a variety of trial states. The first three trial states utilize the same operation that prepares the three clusters, respectively, and are observed to be correctly classified. We further carry out the classification of trial states that are prepared with other operations (See Methods). Such examples are the two states $D(\alpha = 1)(|0\rangle - |1\rangle)$ and $\sqrt{0.3S(r = 0.8e^{i\pi})}(|0\rangle + \sqrt{0.2}(|0\rangle + |1\rangle) + \sqrt{0.5}D(\alpha = 1.7)|0\rangle$ (up to a normalization factor), that are both classified to cluster 2. These results indicate that the outcome of the algorithm is determined by overlap of the quantum states, which can be visualized by the similarity of their Wigner functions. This can be interpreted as a form of Wigner function pattern recognition, much like applying $k$-NN to handwriting recognition in classical ML \cite{19}.

In conclusion, we realize a bosonic learning machine using motional modes in a system of trapped ions. The core subroutine of the machine is the efficient SWAP test that estimates the overlap of two motional modes with a constant gate duration, a feature that is preferred in solving problems with large dimension such as quantum support vector machines. We expect that the experimentally accessible dimensions can be further increased by improving various technical aspects of the setup, such as the motional coherence time and coupling strength between modes. The same bosonic learning machine can also be used to implement several quantum kernels proposed in \cite{17} and on different experimental platforms such as superconducting microwave cavities \cite{21,27}. These results provide a key primitive for developing advanced quantum-enhanced machine learn-
ing algorithms in infinite-dimensional feature spaces.

Acknowledgments

We thank P.T. Le, H.N. Le, and L.C. Kwek for discussions. This research is supported by the National Research Foundation, Prime Minister’s Office, Singapore, and the Ministry of Education, Singapore, under the Research Centers of Excellence program and NRF Quantum Engineering Program (Award QEP-P4)

Author contributions

C.H.N. and D.M. conceived the experiment. C.H.N. K.W.T, G.M. and H.C.J.G. built and maintained the experimental apparatus. C.H.N. and K.W.T performed the experiments and analyzed the experimental data with assistance from G.M. and H.C.J.G. All authors contributed to the manuscript preparation.

Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

Author Information

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.H.N. (ngchihuan@gmail.com) or D.M. (phymd@nus.edu.sg).

Methods

Experimental setup. We trap two $^{171}$Yb$^+$ ions in a four-rod linear radio-frequency Paul trap with single-ion secular trap frequencies $(\omega_x, \omega_y, \omega_z) = 2\pi \times (1.274, 0.945, 0.519)$ MHz. The normal motional modes $A$, $B$, and $C$ are the out-of-phase motion in the $y$ direction, out-of-phase motion and in-phase in the $x$ direction with frequencies $(\omega_A, \omega_B, \omega_C) = 2\pi \times (0.782, 1.159, 1.274)$ MHz, respectively. The transverse trapping frequencies are actively stabilized and drift less than 100 Hz/hour.

One of the ions, pumped to the metastable state $2F_{7/2}$ by driving the $^2D_{3/2} \rightarrow 3[1/2]_{3/2}$ transition at 398.98 nm, does not interact with optical fields and is referred to as the “dark” ion. At the beginning of all experiments, the other ion is Doppler cooled by a 369.53 nm laser red-detuned from the $^2D_{3/2}, F = 1$ to $^2P_{1/2}, F = 0$ transition, and sympathetically cools the “dark” ion. The two hyperfine states used in the experiment are $|g\rangle = |^2S_{1/2}, F = 0, m_F = 0\rangle$ and $|e\rangle = |^2S_{1/2}, F = 1, m_F = 0\rangle$, which are separated by $\omega_0/2\pi = 12.643$ GHz. The ion is initialized via optical pumping in the state $|g\rangle$. Resonance fluorescence state detection technique is used to detect the ion’s internal state [29].

A frequency-doubled, mode-locked Ti:Sapphire laser with a central wavelength of 372.85 nm is used to drive the stimulated Raman transitions. The mode-locked laser has a pulse duration of 3 ps and a repetition rate of 76.2 MHz. Two beams, R1 and R2 (Fig. 3a,b) with an average power of 60 mW each, are sent through acousto-optical modulators (AOMs) and focused to a beam waist of $\sim 15 \mu m$ at the ions. The beams R1 and R2 form $45^\circ$ and $135^\circ$ angles with respect to the $z$-axis (Fig. 3a). The detuning of the Raman beams are controlled by the AOMs.

We begin all experiments with all the radial motional modes prepared in the ground state (residual $n \leq 0.05$) using 10 ms Sisyphus cooling followed by 250 cycles of Raman sideband cooling.

Controlled beam splitter gates. We implement controlled beam splitter gates $U_{CBS}$ between modes $A$ and $B$ by applying a bichromatic Raman beatnote at $\omega_0 \pm \omega_{AB}$, where $\omega_{AB} = |\omega_A - \omega_B|$. The resulting Hamiltonian is

$$H_I = \frac{\hbar g(t)}{2} \sigma_x (a^\dagger e^{-i\psi} + ab e^{i\psi}) = \sigma_x H_{BS},$$

where $\psi = \frac{\pi}{2} - \frac{\omega_0 t}{\hbar}$, and $\psi_C$ and $\psi_B$ are the phases of the bichromatic fields. We let $\sigma_x = (e^{-i\psi_S} \sigma_+ + e^{i\psi_S} \sigma_-)$, where $\psi_S = -\frac{\omega_0 + \omega_x}{\hbar}$, and $\sigma_+ = |e\rangle \langle g|$, $\sigma_- = |g\rangle \langle e|$. $g(t)$ is the coupling strength.

The controlled beam splitter gate corresponds to the unitary evolution $U_{CBS}(\theta, \psi) = \exp[-(i/\hbar) \int_0^T \sigma_x H_{BS}(\psi(t), dt)]$, where $\theta = \int_0^T g(t) dt$. In the experiment, we obtain a coupling strength $g/2\pi \sim 680$ Hz. This corresponds to implementing a beam splitter operation $U_{CBS}(\theta = \pi/2)$ in $T \sim 365 \mu s$, much shorter than the motional phase coherence time for the superposition of $|0\rangle + |1\rangle$ which is approximately 10 ms.

Preparation and characterization of quantum states. The universal embedding circuit relies on a series $n - 1$ pairs of carrier and red-sideband:

$$\prod_{l=1}^{n-1} U_{ab}(\tau_l, \Upsilon_l) U_c(t_l, v_l + v'_l)$$

where $t_l$ (\Upsilon_l) and $v_l$ (\Upsilon_l) are the duration and phase for the $l$-th carrier (red-sideband) pulse that are determined by the Eberly-Law algorithm [24]. The additional phases $v'_l$ compensate for the rotation of the internal states due to the AC Stark shifts of the red-sideband pulses (See Supplementary). The embedding circuit is used to prepare the centroid states of $K$-means, the cluster state $H_{Fock}$ of $K$-means experiment, and the trial state (up to a normalization factor) $\sqrt{0.355(\alpha = 0.8 e^{-i\theta})}|0\rangle + \sqrt{0.7(\alpha = 1.7)|1\rangle}$ of $k$-NN experiment. We use spin-dependent force to prepare all coherent states and the squeezed vacuum states [29]. For the trial state $D(\alpha = 1)|\{0\rangle - |1\rangle\rangle$, we first prepare $|0\rangle - |1\rangle$, followed by a displacement of $\alpha = 1$. 

The real and positive coefficients $x_{j+1}$ of the data states $|d_m⟩$, shown in Fig. 2b, are the square root of population distribution $p(j)$, where $p(j) = |⟨d_m|j⟩|^2$. The population distribution is determined from a Fourier transform of the time evolution while driving the red-sideband transition

$$P_e(\tau) = \sum_{j=1}^{\infty} p(j)e^{-\sqrt{j}\gamma \tau}(1 - \cos(\sqrt{j}\Omega_{\text{rsb}}\tau))/2 \quad (3)$$

where $P_e$ is the probability to detect the ion in the state $|e⟩$, $\Omega_{\text{rsb}}$ is the Rabi frequency of the red-sideband, and $\gamma$ is the decoherence rate of motional states. To determine $p(j)$ a fit constrained to be non-negative and $\sum_{j=0}^{6} p(j) = 1$ is applied to Eq. 3 for $j = 0$ to 6.

[1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature 549, 195 (2017)
[2] X.-D. Cai, D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Physical Review Letters 114, 110504 (2015)
[3] V. Trávníček, K. Bartkiewicz, A. Černoch, and K. Lemr, Physical Review Letters 123, 260501 (2019)
[4] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Nature 567, 209 (2019)
[5] S. Johri, S. Debnath, A. Mocherla, A. Singh, A. Prakash, J. Kim, and I. Kerenidis, Nature 549, 195 (2017)
[6] V. Trávníček, K. Bartkiewicz, A. Černoch, K. Jiráková, K. Lemr, and F. Nori, Scientific Reports 10, 12356 (2020)
[7] T. Xin, L. Che, C. Xi, A. Singh, X. Nie, J. Li, Y. Dong, and D. Lu, Phys. Rev. Lett. 126, 110502 (2021)
[8] E. Peters, J. Caldeira, A. Ho, S. Leichenauer, M. Mohseni, H. Neven, P. Spentzouris, D. Strain, and G. N. Perdue, (2021)
[9] S. L. Braunstein and P. van Loock, Reviews of Modern Physics 77, 513 (2005)
[10] L. Ortiz-Gutiérrez, B. Gabrielly, L. F. Muñoz, K. T. Pereira, J. G. Filgueiras, and A. S. Villar, Optics Communications 397, 166 (2017)
[11] C. Fiumani, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home, Nature 566, 513 (2019)
[12] S. Lloyd, M. Mohseni, and P. Rebentrost, Nature Physics 10, 631 (2014)
[13] S. Lloyd, M. Mohseni, and P. Rebentrost, Nature Physics 10, 631 (2014)
[14] S. Lloyd, G. Plenio, and P. Zanardi, Nature Communications 7, 10138 (2016)
[15] P. Rebentrost, M. Mohseni, and S. Lloyd, Physical Review Letters 113, 130503 (2014)
[16] H.-K. Lau, R. Pooser, G. Siopsis, and C. Weedbrook, Physical Review Letters 118, 080501 (2017)
[17] M. Schuld and N. Killoran, Physical Review Letters 122, 040504 (2019)
[18] S. Lloyd, M. Schuld, A. Ijaz, J. Izaac, and N. Killoran, (2020) arXiv:2001.03622
[19] T. Hastie, R. Tibshirani, and J. H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer series in statistics (Springer, 2009).
[20] R. Filip, Physical Review A 65, 062320 (2002)
[21] Y. Y. Gao, B. J. Lester, Y. Zhang, C. Wang, S. Rosenblum, L. Frunzio, L. Jiang, S. M. Girvin, and R. J. Schoelkopf, Physical Review X 8, 021073 (2018)
[22] H. C. J. Chan, G. Maslennikov, K.-W. Tseng, C. Nguyen, and D. Matsukevich, Phys. Rev. Lett. 124, 170502 (2020)
[23] C.-H. Nguyen, K.-W. Tseng, G. Maslennikov, H. C. J. Gan, and D. Matsukevich, (2021) arXiv:2103.10219
[24] C. K. Law and J. H. Eberly, Physical Review Letters 76, 1055 (1996)
[25] A. Ben-Kish, B. DeMarco, M. Rowe, J. Britton, W. M. Itano, B. M. Jelenkovic, C. Langer, D. Leibfried, T. Rosenband, and D. J. Wineland, Phys. Rev. Lett. 90, 037902 (2003)
[26] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, IEEE Transactions on Information Theory 63, 5628 (2017)
[27] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bianzzi, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, J. Wener, J. M. Martinis, and A. N. Cleland, Nature 459, 546 (2009)
[28] S. Olmschenk, K. C. Younge, D. L. Moehring, D. N. Matsukevich, P. Maunz, and C. Monroe, Physical Review A 76, 052314 (2007)
[29] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, Reviews of Modern Physics 75, 281 (2003).
Supplementary Material: Experimental quantum-enhanced machine learning over infinite dimensions

I. Delays between state preparation and the SWAP test

The two steps required for operation of the bosonic learning machine are the state preparation and the overlap measurement using the SWAP test. However, the quantum state of a harmonic oscillator evolves in time as \( |\Psi\rangle = \sum_j c_j e^{-i\omega j t} |j\rangle \) where \( \omega \) is the motional frequency. If the frequencies of motional modes \( \omega_B \) and \( \omega_C \) are different, care must be taken to calibrate the waiting time between preparation and measurement to ensure that intended states are passed to the SWAP test. We set the delay time as \( nT + \text{offset} \). Otherwise, for example, the overlap of the two coherent states in modes \( B \) and \( C \) will oscillate as a function of the delay between the two steps, with a period of \( T = \frac{2\pi}{|\omega_C - \omega_B|} \).

The trap frequency measurements done by scanning the motional sidebands are affected by the Stark shifts of the internal states. To obtain a frequency difference between the modes which is free from AC Stark shifts we measure the overlap \( |\langle \Psi | \Phi_1 \rangle|^2 \) as a function of the delay time \( \tau \), where \( |\Psi\rangle = |\Phi_1\rangle = \frac{1}{\sqrt{4}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle) \) and \( |\Phi_2\rangle = \frac{1}{\sqrt{4}}(|0\rangle - |1\rangle + |2\rangle - |3\rangle) \). Figure S1 shows the results of overlap measurement. The period \( T \) is obtained from the time difference between two peaks in the \( |\langle \Psi | \Phi_1 \rangle|^2 \) overlap measurement. We estimate \( T \approx 8.2 \mu s \).

II. Additional details on state preparation

The Fourier analysis of the time evolution while driving the red sideband can measure only the diagonal terms (phonon distribution) of the density matrix of the states. To verify that the universal embedding circuit can prepare states with desired off-diagonal terms, we carry out the following experiment. We prepare \( |\Phi_2\rangle = \frac{1}{\sqrt{4}}(|0\rangle - |1\rangle + |2\rangle - |3\rangle) \) and compare the two overlaps \( |\langle \Psi | \Phi_1 \rangle|^2 \) and \( |\langle \Psi | \Phi_2 \rangle|^2 \) as a function of waiting time \( \tau \). Simulations predict that the second overlap \( |\langle \Psi | \Phi_2 \rangle|^2 \) is shifted in time, relative to \( |\langle \Psi | \Phi_1 \rangle|^2 \), by \( T/2 \) if the states with correct off-diagonal terms are prepared. The experimental results shown in Fig. S1 agree with calculations.

FIG. S1: Measurement of \( |\langle \Psi | \Phi_1 \rangle|^2 \) (blue circles) and \( |\langle \Psi | \Phi_2 \rangle|^2 \) (orange squares) as a function of waiting time \( \tau \), where \( |\Psi\rangle = |\Phi_1\rangle = \frac{1}{\sqrt{4}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle) \) and \( |\Phi_2\rangle = \frac{1}{\sqrt{4}}(|0\rangle - |1\rangle + |2\rangle - |3\rangle) \).

FIG. S2: a. Circuit diagram for measuring the AC Stark shift induced by the inserted rsb pulse \( U_{\text{rsb}} \). \( R(\pi/2, \phi) \): \( \pi/2 \) rotation pulses with phase \( \phi \) of the internal states. b. Probability of detecting the internal state in \( |e\rangle \) as a function of the red-sideband pulse duration \( \tau \) with \( \phi = 0 \) (blue circles) and \( \phi = -2\pi \times 5.9 \text{kHz} \) (orange squares). The dashed line is the calculated probability when the AC Stark is compensated.
The red-sideband pulse $U_{\text{rsb}}(\tau_l, \Upsilon_l)$ induces an AC Stark shift $\hbar \Delta \omega_{\text{AC}}$ that gives rise to a rotation $U = e^{-i \Delta \omega_{\text{AC}} \tau_l \sigma_z}$ of the internal states. To compensate for this rotation, an additional phase $\nu_{l+1} = \sum_{l' \leq l} \Delta \omega_{\text{AC}} \tau_{l'}$ is included to the next carrier pulses $U_c(t_l, \nu_l + \nu_{l+1})$. We determine $\Delta \omega_{\text{AC}}$ using the Ramsey experiment with an inserted red-sideband pulse described in Figure S2a. The second $\pi/2$ rotation pulse has a phase $\phi$. We vary the red-sideband pulse duration $\tau$ while monitoring the probability of detecting the internal state in $|e\rangle$. When $\phi$ is set to $-\Delta \omega_{\text{AC}} \tau$ and $\Delta \omega_{\text{AC}} \sim 2\pi \times 5.9$ KHz, the measurement result (Fig. S2b) agrees with the expectation when the AC Stark shift is compensated.

III. EXTENDED DATA

FIG. S3: Step-by-step cluster assignment of $K$-mean algorithm visualized in scatter plot matrix of the centroids (crosses) and data states: $H_{\text{sqz}}$ (circles), $H_{\text{Fock}}$ (squares), and $H_{\text{coh}}$ (diamonds). The error bar denotes the one $\sigma$ fit error of the state characterization.
FIG. S4: The overlap between trial states and all members in each cluster in the \( k \)-nn experiment. The error bars correspond to statistical uncertainty and represent 95% confidence intervals.
FIG. S5: Population distribution of the centroid states in the Fock basis. Blue sideband pulses are used to analyze centroid states of step 0 and red sideband pulses to analyze the rest of the centroid states. Insets show the probability to detect the internal state in $|e\rangle$ as a function of the pulse duration. The red curves are the fits to the population using Eq. 3. The orange bars show the desired population distribution.