Realizing quantum linear regression with auxiliary qumodes

Dan-Bo Zhang,[1] Zheng-Yuan Xue,[1] Shi-Liang Zhu,[2,1,*] and Z. D. Wang[3,†]
[1]*Guangdong Provincial Key Laboratory of Quantum Engineering and Quantum Materials, and School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China*
[2]*National Laboratory of Solid State Microstructures, School of Physics, Nanjing University, Nanjing 210093, China*
[3]*Department of Physics and Center of Theoretical and Computational Physics, University of Hong Kong, Pokfulam Road, Hong Kong, China*

(Received 5 October 2018; revised manuscript received 27 December 2018; published 18 January 2019)

In order to exploit quantum advantages, quantum algorithms are indispensable for operating machine learning with quantum computers. We here propose an intriguing hybrid approach of quantum information processing for quantum linear regression, which utilizes both discrete and continuous quantum variables, in contrast to existing wisdoms based solely upon discrete qubits. In our framework, data information is encoded in a qubit system, while information processing is tackled using auxiliary continuous qumodes via qubit-qumode interactions. Moreover, it is also elaborated that finite squeezing is quite helpful for efficiently running the quantum algorithms in a realistic setup. Comparing with an all-qubit approach, the present hybrid approach is more efficient and feasible for implementing quantum algorithms, still retaining exponential quantum speed-up.

DOI: 10.1103/PhysRevA.99.012331

I. INTRODUCTION

For quantum systems, there exist both discrete and continuous variables, and they may interact with each other. A typical example is an light-atom interacting system, which demonstrates how discrete energy levels and continuous light fields evolve quantum mechanically. Notably, a hybrid scenario of information processing using both discrete and continuous variables has been proposed before [1–5], such as hybrid quantum computing [3–5]. The relevant advantages have been indicated in tasks like quantum float computing [6] and quantum phase estimation (QPE) [3,7], in which infinite dimensions of continuous variables are exploited [7–10], making the proposals promising.

As is well known, machine learning plays an important role for extracting worthwhile information and making trustable predictions in an era of big data [11]. In addition, a magnificent combination of machine learning and quantum mechanics has opened a new window for information processing [12–19]. A class of quantum machine learning [19–22] is based on the Harrow-Hassidim-Lloyd (HHL) algorithm, which aims to obtain the inverse of a matrix with exponential speed-up under reasonable conditions [20]. Normally, quantum linear regression has been regarded as a representative task in quantum machine learning and investigated by various all-qubit approaches [13,19,23]. The algorithms for quantum linear regression usually consist of several main parts, namely, quantum phase estimation, regularization, singular value transformation, and prediction. All of these approaches require ancillary qubits to register singular values in quantum phase estimations that are necessary for quantum linear regression. Depending on the desired precision, the number of ancillary qubits may be large, which demands a rarer resource of qubits and is a grand challenge for current quantum technology. On the other hand, the quantum phase estimation can be implemented much more efficiently using the hybrid approach [3], which requires merely one qumode based on appropriate squeezing states [7]. This motivates us to work out a quantum algorithm of linear regression using the hybrid approach with desired quantum advantages.

In this paper, we first introduce how to convert a linear regression into a quantum task. We then recall the quantum algorithms of all-qubit systems and analyze their properties. We emphasize the demanding of ancillary qubits and the inefficiency of incorporating regularization and realizing singular value transformation. By introducing qumodes, we propose a hybrid approach for quantum linear regression, where single values are encoded into the entangled two qumodes and single-value transformation can be implemented simply by homodyne measurements with postselection. This makes the algorithm more feasible for future physical implementation as it involves basic continuous variable quantum operations, in contrast to the all-qubit approach that requires complicated quantum circuits for quantum arithmetic and control rotation. A brief proposal for a physical realization of the algorithm with trapped ions is suggested. Our results show that the hybrid approach still retains the same order of runtime $O(\log MN)$ as the all-qubit approach. The regularization can be incorporated by a controlled-phase gate on two qumodes, and it may greatly reduce the required squeezing factor for the case of a bad condition number [10,20]. We also investigate a basic role of the finite squeezing factor and find that it is not only helpful for running the algorithm efficiently, but also may be taken as an extra regularization for regression.

The paper is organized as follows. We introduce quantum linear regression in Sec. II and give an analysis of the existing all-qubit approach in Sec. III. Then in Sec. IV we present the quantum algorithm using qumodes and its physical

*slzhu@nju.edu.cn
†zwang@hku.hk

implementation in trapped ions. Finally, discussions and conclusions are given in Sec. V.

II. QUANTUM VERSION OF LINEAR REGRESSION

In this section we will demonstrate how to formulate the linear regression as a quantum problem that may be solved using a quantum algorithm. Let us first introduce the linear regression. Given a training data set of M points $\{\mathbf{a}^{(m)}, y^{(m)}\}$, where $\mathbf{a}^{(m)} \in R^N$ is a vector of N features and $y^{(m)} \in R$ is the target value, the goal is to learn a linear model with parameters $\mathbf{w} \in R^N$ that can give the prediction $\tilde{y}$ for new data $\tilde{\mathbf{a}}$ as $\tilde{y} = \tilde{\mathbf{a}}^T\mathbf{w}$. The parameters $\mathbf{w}$ can be estimated by minimizing the loss function of least-square errors

$$L(\mathbf{w}) = \sum_{m=1}^{M} \left(\mathbf{w}^T\mathbf{a}^{(m)} - y^{(m)}\right)^2 + \chi||\mathbf{w}||^2 \tag{1}$$

over $\mathbf{w}$. Here $\chi||\mathbf{w}||^2$ is a regularization term with parameter χ, which is usually considered in machine learning for better performance. It will be shown later that the regularization is helpful for efficiently implementing the quantum algorithm. Introducing the matrix notation $A_{mn} = \mathbf{a}_n^{(m)}$, the linear regression solution turns out to be $\mathbf{w} = A^+\mathbf{y}$, where the Moore-Penrose pseudoinverse (with χ term) reads $A^+ = (A^TA + \chi I)^{-1}A^T$. Here I is the identity matrix. It is inspiring to study the linear regression by using the singular value decomposition of A from Ref. [23], $A = \sum_i \lambda_i \mathbf{u}_i\mathbf{v}_i^T$. Here λ_i are singular values of A with corresponding left (right) eigenvector $\mathbf{v}_i$ ($\mathbf{u}_i$). Now it can be verified that the Moore-Penrose pseudoinverse reads as $A^+ = \sum_i \frac{\lambda_i}{\lambda_i^2+\chi}\mathbf{v}_i\mathbf{u}_i^T$. Then the prediction can be written as $\tilde{y} = \sum_i \frac{\lambda_i}{\lambda_i^2+\chi}\mathbf{u}_i^T\mathbf{y}\tilde{\mathbf{a}}^T\mathbf{v}_i$.

To formulate a quantum version of linear regression, we need to encode each component as a quantum state or a quantum operator. For a vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, the amplitude encoding scheme is taken to encode $\mathbf{x}$ into a quantum state as $|\psi_\mathbf{x}\rangle = \sum_n x_n|n\rangle$. Here $|n\rangle = |n_1n_2n_3\cdots\rangle$,where $n = n_1n_2n_3\cdots$ is a binary representation of integer n, and it requires a number of $[\log_2 N]$ qubits. Accordingly, $\mathbf{y}$ and $\tilde{\mathbf{a}}$ are encoded as $|\psi_\mathbf{y}\rangle = \sum_m y^{(m)}|m\rangle$ and $|\psi_{\tilde{\mathbf{a}}}\rangle = \sum_n \tilde{\mathbf{a}}_n|n\rangle$, respectively. Without loss of generality, we assume all quantum states in this paper are normalized, and the final result should be rescaled accordingly. The remaining question now is how to treat A^+. It is natural to take it as a operator, and the prediction finally writes as $\tilde{y} = \langle\psi_{\tilde{\mathbf{a}}}|A^+|\psi_\mathbf{y}\rangle$. One should pay special attention to the following aspects for this approach [13]. First, A^T is not necessarily a square matrix, and the Hilbert space should be extended to define a square matrix that contains A^T. Second, two HHL-like algorithms for applying A^T and $(A^TA + \chi I)^{-1}$ sequentially are required. Third, both A^T and $(A^TA + \chi I)^{-1}$ are not unitary in general, there must be some nonunitary procedure such as measurements or projections in the algorithm. Thus, measurement is required in the middle of the algorithm.

An alternative approach to circumvent the above problems is to treat A^+ as a quantum state [23]. This is possible by rewriting the prediction as a tensor formula

$$\tilde{y} = \sum_i \frac{\lambda_i}{\lambda_i^2+\chi}(\mathbf{u}_i \otimes \mathbf{v}_i)^T\mathbf{y} \otimes \tilde{\mathbf{a}},$$

which is an inner product between two vectors. Accordingly, we may write the quantum state that corresponds to A^+ as $|\psi_{A^+}\rangle = \sum_i \frac{c\lambda_i}{\lambda_i^2+\chi}|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle$, where c is introduced for normalization. Then the prediction $\tilde{y}$ is obtained as the inner product between $|\psi_{A^+}\rangle$ and $|\psi_\mathbf{y}\rangle \otimes |\psi_{\tilde{\mathbf{a}}}\rangle$, up to a constant factor $\frac{1}{c}$.

On the other hand, the input data A can be loaded as a quantum state $|\psi_A\rangle = \sum_{m,n} \mathbf{a}_n^{(m)}|m\rangle|n\rangle = \sum_i \lambda_i|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle$, where the second equality comes from Schmidt decomposition. Thus, the key task is to find a quantum algorithm that transforms $|\psi_A\rangle$ to $|\psi_{A^+}\rangle$. After loading data A as the initial state $|\psi_A\rangle$, different quantum techniques may be used to steer the system from initial state $|\psi_A\rangle$ to the target state $|\psi_{A^+}\rangle$.

Since both $|\psi_A\rangle$ and $|\psi_{A^+}\rangle$ are bipartite quantum states, an investigation of their entanglement structure may inspire the design of the quantum algorithm. They share the same entangling basis of $\{|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle\}$ with different coefficients, where $|\psi_{\mathbf{v}_i}\rangle \in R^N$ are eigenstates in feature space and $|\psi_{\mathbf{u}_i}\rangle \in R^M$ are eigenstates related to sample space. Then a quantum algorithm can be devised that keeps those basis unchanged, while the coefficients are transformed in the form $g(\lambda) = \frac{\lambda}{\lambda^2+\chi}$. Such a transformation has been realized using ancillary qubits [13,20,23].

It should be pointed out that focusing on $|\psi_{A^+}\rangle$ instead of the states of parameters $|\psi_\mathbf{w}\rangle$, which is defined as $|\psi_\mathbf{w}\rangle = \sum_n w_n|n\rangle$ for parameters $\mathbf{w} = (w_1, \ldots, w_n)$, has its own advantages, especially when there are multiple target values for linear regression. For example, based on the features of one person, one can predict two target values such as both income and cost. In such a case, predictions are obtained by inner product of $|\psi_{A^+}\rangle$ and $|\psi_{\tilde{\mathbf{a}}}\rangle|\psi_{\mathbf{y}_l}\rangle$ ($l = 1, 2$ for income and cost), separately. In other words, once $|\psi_{A^+}\rangle$ is obtained it can be applied for linear regression with different target vectors. On the other hand, $|\psi_{\mathbf{w}_t}\rangle$ is specified to fixed target value. Moreover, one may construct the later from the former for each target vector $|\psi_{\mathbf{y}_t}\rangle$ as

$$|\psi_{\mathbf{w}_t}\rangle = \sum_i \frac{c\lambda_i}{\lambda_i^2+\chi}\langle\psi_{\mathbf{y}_t}|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle. \tag{2}$$

In this sense, $|\psi_{A^+}\rangle$ is more fundamental than $|\psi_{\mathbf{w}_t}\rangle$. In our implementation, we adopt this more efficient way.

III. EXISTING ALL-QUBIT APPROACH

Before proposing the hybrid approach using both qubits and qumodes, we first recall how to do quantum linear regression with all-qubit systems, where regularization has not been considered [13,19,23]. An analysis is given to show that the all-qubit approach demands many ancillary qubits in several main procedures of quantum linear regression. This motivates us to propose a hybrid approach that is more efficient.

Basically, there are two subroutines in the algorithm of the all-qubit approach. First, quantum phase estimation is applied that registers λ_i in ancillary qubits; second, coefficients $\frac{1}{\lambda_i}$ are obtained by a conditional rotation on an extra ancillary qubit and then a projection to the $|1\rangle$ state. The success rate of projecting to $|1\rangle$ is proportional to the condition number $\kappa = \frac{\lambda_{\max}^2}{\lambda_{\min}^2}$, where $\lambda_{\max/\min}$ stands for largest or smallest singular values of A. Under a bad condition number, e.g., $\kappa_c \sim O(N)$

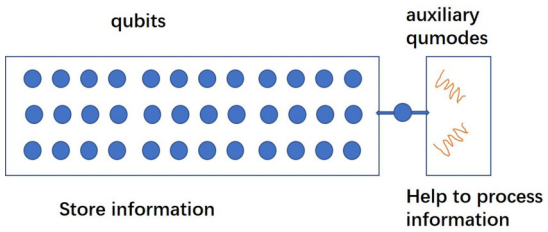

FIG. 1. Illustration of the hybrid approach that exploits the best of both qubits and qumodes. In our implementation of quantum linear regression, qubits and qumodes are minimally coupled by connecting only to an ancillary qubit.

(which naturally arises for low-rank A^TA), the runtime will scale with $O(N)$ and thus destroy the exponentially speed-up. This issue, as will be demonstrated later, can be remedied by introducing regularization.

We analyze the cost of resources, especially on the number of qubits. For the quantum phase estimation, first, a series of $\{e^{iA^TA\tau}\}$ with different time τ should be constructed. Second, since λ_i are typically continuous numbers, to encode them with precision ε a number of $O(\log \varepsilon^{-1})$ qubits is required. If regularization is included, a quantum circuit for quantum addition is required, which requires extra $O(\log \varepsilon^{-1})$ qubits. For singular value transformation, to realize the required conditional rotation an oracle should be constructed which would involve many qubits depending on the precision. Importantly, for all processes, although many ancillary qubits should be used, they would be discarded finally. Since their role is just to register singular values and further give singular value transformation, it is desirable to have an alternative approach that requires less resources of qubits.

It is known that the quantum phase estimation as well as the quantum addition are more naturally and efficiently implemented in a quantum system with continuous variables [6] (also see Appendix A). On the other hand, vectors of classical data are better encoded in quantum states of qubit systems, as each component of a vector now becomes the corresponding amplitude directly, while encoding in qumodes this would be much complicated [8,9]. Given the above two considerations, it naturally calls for a hybrid quantum computing approach that exploits both the advantages of qubits and qumodes (see, e.g., Fig. 1), as will be explored in this paper.

IV. QUANTUM ALGORITHM USING TWO AUXILIARY QUMODES

We present the quantum algorithm for linear regression using two auxiliary qumodes. Further discussions will be presented in Sec. V. We introduce qumodes $|x\rangle_q$ and $|x\rangle_p$, which are eigenstates of conjugate quadrature field operators $\hat{q}$ and $\hat{p}$ (such as momentum and position operators), respectively. As $[\hat{q},\hat{p}]=i$, we have $|q\rangle_q=\frac{1}{\sqrt{2\pi}}\int dp e^{-iqp}|p\rangle_p$. It should be emphasized that the squeezing factor cannot be infinite physically. Nevertheless, we first present the algorithm for the case of infinite squeezing, as the main procedures of the algorithm are more clearly revealed. We then stress the importance of finite squeezing and modify the algorithm accordingly.

The algorithm under infinite squeezing can be summarized as following:

(1) *State preparation*. Load training data $\{\mathbf{a}^{(m)}\}$ into a quantum state $|\psi_A\rangle=\sum_m|m\rangle|\psi_{\mathbf{a}^{(m)}}\rangle$ with quantum random access memory [24–26]. Two qumodes are initialized in $|0\rangle_{q_1}|0\rangle_{q_2}$.

(2) *Quantum phase estimation*. Perform $U=e^{i\eta A^TA\hat{p}_1\hat{p}_2}$ on $|\psi_A\rangle|0\rangle_{q_1}|0\rangle_{q_2}$ which leads to

$$\sum_i \lambda_i|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle\int dp|p\rangle_{p_1}\left|\eta\lambda_i^2 p\right\rangle_{q_2}.$$

This step encodes singular values of A^TA into the entangled two qumodes.

(3) *Regularization*. For a preset regularization parameter χ which is a hyperparameter, perform $e^{i\eta\chi\hat{p}_1\hat{p}_2}$ that shifts the q_2 mode to $|\eta(\lambda_i^2+\chi)p\rangle_{q_2}$. Different values of χ can be tried that a specified χ can be chosen to get a good enough performance of prediction.

(4) *Singular-value transformation*. After a homodyne detection on qumode q_2 with the result $q_2=0$, the state collapses to $|\psi_{A^+}\rangle$, where qumodes are discarded since they are disentangled from qubits. This step transforms quantum amplitudes from $\{\lambda_i\}$ to $\{\frac{\lambda_i}{\lambda_i^2+\chi}\}$.

(5) *Prediction*. For new data $\tilde{\mathbf{a}}$, prepare the reference state $|\Psi_R\rangle=|\mathbf{y}\rangle|\psi_{\tilde{\mathbf{a}}}\rangle$. The prediction $\tilde{y}$ is proportional to the inner product between the target state $|\psi_{\mathbf{A}^+}\rangle$ and $|\Psi_R\rangle$, denoted as $\tilde{y}'=\langle\psi_{\mathbf{A}^+}|\Psi_R\rangle$, which can be implemented using the method introduced in Refs. [16,27]. The key point is to introduce an ancillary qubit to first construct an entangled state $|\Psi\rangle=\frac{1}{\sqrt{2}}(|0\rangle|\Psi_R\rangle+|1\rangle|\psi_A\rangle)$. This can be achieved by setting the ancillary qubit in $(|0\rangle+|1\rangle)/\sqrt{2}$ as a control state, and when the ancillary qubit is in $|0\rangle$ the reference state $|\Psi_R\rangle$ is prepared, and when in $|1\rangle$ the state $|\psi_A\rangle$ is prepared. Then, conditioned on the $|1\rangle$ state, procedures of $1-4$ for the algorithm are performed to generate $|\psi_{A^+}\rangle$, e.g., by attaching a control to all operations in procedures 1–4, which lead to a state $|\Psi'\rangle=\frac{1}{\sqrt{2}}(|0\rangle|\Psi_R\rangle+|1\rangle|\psi_{A^+}\rangle)$. Then a σ_x measurement on the ancillary qubit of quantum state $|\Psi'\rangle$ gives the output $+1$ with probability $p=\frac{1}{2}(1+c\tilde{y}')$. Thus $\tilde{y}'=\frac{2p-1}{c}$.

The above-presumed infinite squeezing factor is impractical in implementation, and the success rate for postselecting $q_s=0$ is vanishing.

We now take finite squeezing into account, which leads to a modification of the algorithm in three places. First, in the state preparation two qumodes are prepared as

$$|G_{12}\rangle=s^{-1}\pi^{-\frac{1}{2}}\int dp_1\,dp_2 e^{-(p_1^2+p_2^2)/2s^2}|p_1\rangle_{p_1}|p_2\rangle_{p_2}$$

with a squeezing factor s. Since $|q\rangle_q=\frac{1}{\sqrt{2\pi}}\int dp e^{-iqp}|p\rangle_p$, $|G_{12}\rangle=|0\rangle_{q_1}|0\rangle_{q_2}$ at $s\to\infty$ limit. Second, in the singular-value transformation, to fully disentangle qubits and qumodes, two qumodes are postselected as $q_1=Q_1$ and $q_2=Q_2$ using homodyne detection. To get a nonzero success rate, we can let the point (Q_1,Q_2) locate within a small area proportional to ϵ_q centered at (0,0) (see Appendix B). Now the unnormalized

state can be written as

$$|\psi_{A'}\rangle = \sum_i f_i(Q_1, Q_2)\frac{\lambda_i}{\lambda_i^2+\chi}|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle, \quad (3)$$

where $f_i(Q_1, Q_2) \sim e^{-(Q_1^2+Q_2^2)/2\alpha_i^2 s^2}$ at the limit $\alpha_i^2 s^4 \sim \varepsilon_q^{-1}$. Here $\alpha_i = \eta(\lambda_i^2+\chi)$. Compared with $|\psi_{A^+}\rangle$, each coefficient now is multiplied by $f_i(Q_1, Q_2)$ correspondingly. When $\epsilon_q \to 0$, the state $|\psi_{A'}\rangle$ reduces to $|\psi_{A^+}\rangle$. In this sense ϵ_q characterizes an error. We may take $f_i(Q_1, Q_2)$ as an extra regularization due to finite squeezing, which will be discussed in Sec. V. Third, we use $|\psi_{A'}\rangle$ in the prediction. Note $\tilde{y}'$ is up to a factor to the required prediction $\tilde{y}$ as $\tilde{y} = c'\tilde{y}'$. Such a factor can be obtained as the ratio between $y^{(m)}$ and $\tilde{y}^{(m)}$. Here $\tilde{y}^{(m)}$ is the predicted value of training data $\mathbf{x}^{(m)}$. The factor c' can be more accurate by averaging on predicted values of a number of training data.

At this stage, we proceed to give some implementation details. While involving only qumodes, standard techniques of Gaussian quantum information processing [28] are sufficient for the algorithm, which includes the following procedures. In the state preparation, the squeezing state $|G_{12}\rangle$ can be obtained using a displacement operator and squeezing operator. In the regularization step the operator is just a conditioned-phase gate. Finally, in the singular value transformation, homodyne detection of q modes is required.

The essential part of the implementation of the algorithm is to construct the operator $U = e^{i\eta A^T A\hat{p}_1\hat{p}_2}$ for the quantum phase estimation. As U extracts singular values implied in states of qubits and registers them on entangled qumodes, it is essentially a hybrid quantum operator. Effectively constructing U involves a density matrix exponentization of $A^T A$ using the method in Refs. [15,29]. Further, the required exponential swap gate should be modified [9] to couple the density matrix $A^T A$ to $\hat{p}_1\hat{p}_2$. The density matrix exponentization $e^{i\rho t}$ of density $\rho = A^T A$ can be obtained by repeatedly applying the exponential swap operation and tracking out ρ,

$$\mathrm{Tr}_\rho(e^{i\delta t S_m}\rho\otimes\rho' e^{-i\delta t S_m}) = e^{i\delta t\rho}\rho' e^{-i\delta t\rho} + O(\delta t^2). \quad (4)$$

Here S_m is the swap operator.

To construct U, the exponential swap operator should be modified as (see Fig. 2)

$$e^{i\delta t\eta\hat{p}_1\hat{p}_2 S_m} = C_S R_a^x(\delta t\eta\hat{p}_1\hat{p}_2)C_S. \quad (5)$$

Here $R_a^x(\delta t\eta\hat{p}_1\hat{p}_2) = e^{i\delta t\eta\hat{p}_1\hat{p}_2\sigma_x}$, and $C_S = C_S^{bb'}C_S^{dd'}\cdots$ is a control swap operator on multimodes, where $C_S^{\alpha\alpha'}$ swap between α and α', conditioned on the ancillary qubit a. Note that Eq. (5) is a hybrid approach version of the one in an optical system [9]. The conditional swap operator is constructed as $C_S^{bb'} = C_{bb'}T_{ab'b}C_{bb'}$, where $C_{bb'}$ is a Control-NOT gate that takes $b(b')$ as the control (target), and $T_{ab'b}$ is the Toffoli gate that is conditioned on both b' and the ancillary qubit a. Initially the ancillary qubit a is set in state $|+\rangle_a = \frac{1}{\sqrt{2}}(|0\rangle_a + |1\rangle_a)$. We now have

$$\begin{aligned}&\mathrm{Tr}_\rho(e^{i\delta t\eta\hat{p}_1\hat{p}_2 S_m}\rho\otimes\rho' e^{-i\delta t\eta\hat{p}_1\hat{p}_2 S_m})\\ &= 1 + i[\rho\hat{p}_1\hat{p}_2, \rho']\eta\delta t + O(\delta t^2)\\ &= e^{i\delta t\eta\rho\hat{p}_1\hat{p}_2}\rho' e^{-i\delta t\eta\rho\hat{p}_1\hat{p}_2} + O(\delta t^2).\end{aligned}$$

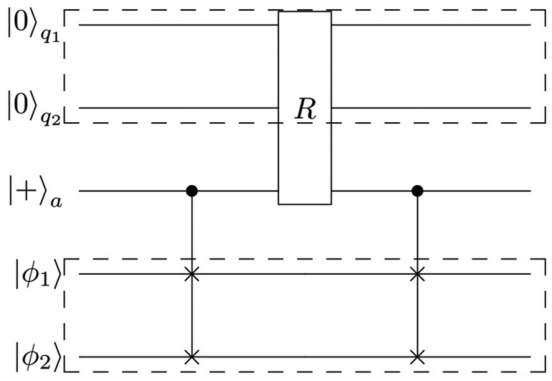

FIG. 2. Quantum circuit for implementing a modified exponential swap gate. An ancillary qubit is prepared at $|+\rangle$. Qumodes and qubits are not directly coupled.

Setting $\rho = A^T A$ we obtain $e^{i\delta t\eta A^T A\hat{p}_1\hat{p}_2}$. To construct U with desired precision ε requires $O(\varepsilon^{-1})$ copies of ρ [8].

A key component of our hybrid approach is to construct the hybrid operator $R_a^x(\delta t\eta\hat{p}_1\hat{p}_2) = e^{i\delta t\sigma_x\hat{p}_1\hat{p}_2}$, which involves a coupling of one qubit to two qumodes. Such a term would not appear naturally in a physical system. Nevertheless, it can be obtained by using basic hybrid quantum operators [3], with the following intertwined quantum evolution: $e^{iH_2\delta t}e^{iH_1\delta t}e^{-iH_2\delta t}e^{-iH_1\delta t} = e^{-[H_1,H_2]\delta t^2} + O(\delta t^3)$. Let $H_1 = g\hat{p}_1\sigma_y$ and $H_2 = g\hat{p}_2\sigma_z$ here and the evolution repeat $O(\frac{1}{g^2\delta t})$ times, and we thus have $R_a(\delta t\eta\hat{p}_1\hat{p}_2)$.

We briefly discuss how to implement the algorithm physically in trapped ions, with some necessary modification. A more detailed demonstration can be found in Appendix C. It should be emphasized that the hybrid approach is more feasible for a physical implementation, compared to the all-qubit approach. For trapped ions, the internal states and motional states of trapped ions can respectively serve as discrete and continuous variables for encoding information, and notably both states are well controllable in the trapped-ion systems [30–35]. The setup contains ions in a Pauli trap lined in the z direction. A specified ion, called an a-ion, provides its motion modes in x and y directions as two qumodes, and one qubit as the ancillary qubit. We initialize two qumodes in squeezing states of momentum and the ancillary qubit in $|0\rangle_a$. Also, n copies of $|\psi_A\rangle$ and the reference state $|\psi_R\rangle$ are prepared using internal states of ions. In the phase estimation, the conditional exponential swap operator can also be written as $e^{i\delta t\eta\hat{q}_x\hat{q}_y S_m} = C_S H_a R_a^z(\delta t\eta\hat{q}_x\hat{q}_y)H_a C_S$ for convenience in the present system of trapped ions, where H_a is the Hadamard gate on the ancillary qubit. The conditional swap operator C_S has been realized experimentally in trapped ions [36]. Setting $H_1(\hat{q}_x) = g\hat{q}_x\sigma_x$, $H_2(\hat{q}_y) = g\hat{q}_y\sigma_y$, which can be realized using lasers for simulating Dirac equations [34,35], the hybrid gate $R_a^z(\delta t\eta\hat{q}_x\hat{q}_y)$ is realizable, using the intertwined quantum evolution of H_1 and H_2. The regularization demands for a controlled-phase gate $C_z = e^{i\chi\eta\hat{q}_x\hat{q}_y}$. Note that $e^{ik_0 t(\hat{q}_x+\hat{q}_y)^2}e^{-ik_0 t\hat{q}_x^2}e^{-ik_0 t\hat{q}_y^2} = e^{i2k_0 t\hat{q}_x\hat{q}_y}$. Then by relaxing the trap along x, y directions and strengthening it along the $x = -y$ direction for the a-ion, we can realize C_z approximately by setting $2k_0 t = \chi\eta$. The singular value transformation can be implemented with homodyne detection of approximately

zero momentum, using the motion tomography method in Ref. [37]. In the prediction, a swap test [38] can access $|\tilde{y}| = \sqrt{2p-1}$, where $p = \frac{1}{2}(1 + |\langle\psi_{A^+}|\psi_R\rangle|^2)$ is the success rate of projecting the ancillary qubit into the $|0\rangle_a$ state. The sign of $\tilde{y}$ should be determined by other means [23].

V. DISCUSSION AND CONCLUSIONS

The hybrid approach for quantum linear regression is similar to that of the all-qubit approach. Both use auxiliary modes to encode information of singular values and finally project out the auxiliary modes, which at the same time gives the singular value transformation. However, the hybrid approach exploits the infinite dimension of qumodes and is essentially different in several aspects. First, it uses entangled qumodes to achieve the singular value transformation. This is unique for continuous variable quantum states (see Appendix A for a more general investigation). Second, the finite squeezing factor s plays a basic role which is absent for the all-qubit approach. Those lead to some distinct properties for the algorithm, as will be revealed below.

We first analyze the runtime of the hybrid approach using two qumodes. It mainly arises from the construction of the operator $U = e^{i\eta A^T A \hat{p}_1 \hat{p}_2}$ for the quantum phase estimation and the success rate of homodyne detection at the step of singular value transformation. To construct U requires $O(\varepsilon^{-1})$ copies of density matrix $A^T A$ with desired precision ε. Meanwhile, each copy of density matrix $A^T A$ can be obtained from $|\psi_A\rangle$ that takes runtime $O[\log(MN)]$ using qRAM. Thus the runtime in the phase estimation is $O[\varepsilon^{-1}\log(MN)]$. The runtime for singular value transformation turns to be $O(\varepsilon_q^{-\frac{3}{2}})$ (see Appendix B). In total, the runtime scales as $O[\varepsilon_q^{-\frac{3}{2}}\epsilon^{-1}\log(MN)]$.

We investigate the role of the regularization parameter χ and the squeezing factor s for quantum speed-up and accuracy. For low-rank $A^T A$, we expect $\lambda_{\min} \sim O(1)$ and $\lambda_{\min} \sim O(N^{-\frac{1}{2}})$, which corresponds to the case of a bad condition number. Without regularization, the exponential speed-up loses in the all-qubit approach [16]. In our hybrid approach, however, exponential speed-up always holds under a constant error ϵ_q once we set $s \sim O(N^{\frac{1}{2}})$ [recalling that $(\lambda_i^2 + \chi)^2 s^4 \sim \epsilon_q^{-1}$]. Moreover, with regularization, the required squeezing factor becomes $s \sim O[\sqrt{\log(N)}]$ if we set $\chi \sim O(\log N)$, which greatly reduces the requiring resources of squeezing. For a preset constant χ, the requiring squeezing factor turns to be $s \propto \frac{1}{\sqrt{\chi}}$.

For the aspect of an experimental reachable squeezing factor, the squeezing of motional modes of trapped ions can reach about 12.6 dB [39] (a squeezing factor about $s \sim 18$), corresponding to four qubits for encoding a continuous number. The precision can be enough for our demonstration proposal of quantum linear regression with trapped ions where only two singular values need to be distinguished (see Appendix C where we have used a data set with four samples and two features). The squeezing factor can be even higher, which may be raised, e.g., with a longer parametric drive duration time when keeping coherence [40]. We may see that raising the squeezing factor requires an improvement of operations on a single qumode against the noisy environment [39,40]. Such a challenge is different from that of increasing the number of qubits, which depend on the scalability of physical setups.

Due to the regularization and finite squeezing, the obtained state $|\psi'\rangle$ is different from one with infinite squeezing and without regularization. To characterize this difference, we calculate the fidelity between $|\psi'\rangle$ and $|\psi_A^+\rangle$ and study its relation to the regularization factor χ and squeezing factor s. More details can be found in Appendix D. As expected, the fidelity increases when s increases as one should expect, and decreases when χ increases. Moreover, under larger regularization the fidelity can decrease significantly more slowly with reducing squeezing factor s.

We further mention the shrinkage effects [41] due to the regularization and finite squeezing, which may benefit machine learning. First, regularization introduces a shrinkage that the coefficients in $|\psi_{A^+}\rangle$ are $\frac{\lambda_i}{\lambda_i^2+\chi}$, instead of $\frac{1}{\lambda}$ for the case without regularization. The extent of shrinkage is higher for small singular values (low-variance components) and lower for large singular values (high-variance components). Further, in the case of finite squeezing, the final state we obtain is $|\psi_{A'}\rangle$ instead of $|\psi_{A^+}\rangle$, that is, each coefficient is rescaled by $f_i(Q_1, Q_2)$, correspondingly. This further reduces the weighting of low-variance components. Rather than taking $f_i(Q_1, Q_2)$ as an imperfection of the algorithm, we can consider it as an extra regularization due to finite squeezing. Whether this regularization will improve the performance of prediction is an interesting question, and we will leave it for further investigation.

It is meaningful to highlight the role of squeezing factor s in the algorithm. On one hand, it is related to the success rate of postselection. The smaller s, the bigger ϵ_q, and thus the higher the success rate. On the other hand, smaller s leads to a bigger shrinkage. Thus, a proper squeezing factor s not only is necessary for efficiently implementing the quantum algorithm but also may be beneficial for machine learning.

In summary, we have adopted a hybrid approach for quantum linear regression, exploiting the advantage of qubits for encoding data and the efficiency of auxiliary qumodes for implementing phase estimations and singular value transformation. The regularization can be incorporated directly with a controlled phase gate on two qumodes. This regularization can remedy the issue of bad condition number by reducing the requirement of squeezing resource and thus is important for running the quantum algorithm efficiently. Our hybrid approach has the same order of runtime as the all-qubit regarding the dimension of data N, the number of training data (samples) M, and the precision ε, but can save using ancillary qubits at the price of requiring finite squeezing states with qumodes. We have also demonstrated the important role of finite squeezing for efficiently running the algorithm. Moreover, we have shown that the shrinkage effect due to finite squeezing may provide a new type of regularization for linear regression. We hope that the hybrid approach may allow us to design quantum algorithms with more flexibility and efficiency, and may even provide new insights for quantum machine learning, by exploiting the infinite dimensionality as well as the finite squeezing nature of qumodes.

ACKNOWLEDGMENTS

We thank Mile Gu for helpful discussions. This work was supported by the NKRDP of China (Grant No. 2016YFA0301800) and the NSFC (Grants No. 91636218 and No. 11474153) as well as the KPST of Guangzhou (Grant No. 201804020055).

APPENDIX A: QUMODES FOR QUANTUM COMPUTING

Quantum phase estimation. At the heart of QPE is quantum Fourier transformation (QFT), which gives exponential speed-up, as performing QFT on N qubits takes the order of N^2 quantum operations, while fast Fourier transformation takes $O(N2^N)$. It is noted that Fourier transformation for quadrature field operators $\hat{p}$ and $\hat{q}$ is inherent, as revealed by the relation of their eigenstates $|q\rangle$ and $|p\rangle$, which takes the form $|q\rangle = \frac{1}{\sqrt{2\pi}} \int dp e^{-iqp}|p\rangle$. Based on the inherent QFT in qumodes, a quantum phase estimation [3] routine can be much simplified. Note that $H|e_i\rangle = E_i|e_i\rangle$. To write eigenvalues into the registering qumode, one just needs to perform a quantum gate $U = e^{iH\hat{p}}$ on $|\psi\rangle|0\rangle_q$, which results in $\sum_i \psi_i|e_i\rangle|E_i\rangle_q$. The registering of eigenvalues in qumodes is just a shift of position (momentum), with the quantity determined by the eigenvalues obtained as H works on its eigenstate. Remarkably, only a single qumode with squeezing factor [7] ε^{-1} is required for desired precision ε. In contrast, for all-qubit system, registering E_i with the same precision requires $O(\log \varepsilon^{-1})$ qubits.

Quantum addition. We consider the addition $a + b$. First encode a as $|a\rangle_q$, then the addition can be realized by performing the shift operator $e^{ib\hat{p}}$ on $|a\rangle_q$, and it is easy to see that the output is $|a + b\rangle_q$. In general, floating computing is more conveniently implemented in the continuous-variable quantum computing setup [6].

Transformation of quantum amplitudes using entangled qumodes. The task is to transform the initial state $\sum_i |\psi_i\rangle$ to $\sum_i g(\lambda_i)|\psi_i\rangle$ (unnormalized), where λ_i relates to $|\psi_i\rangle$, such as a pair of eigenvalue and eigenstate. Here quantum amplitudes are mapped as $a_i \to a_i g(\lambda_i)$. We show how this can be achieved using entangled qumodes and projection. This is unique for entangled states of a continuous variable. For illustration we give the derivation only for infinite squeezing. We assume that the following quantum state can be prepared:

$$|\psi\rangle = \sum_i |\psi_i\rangle \int dp |p\rangle_{p_1} |\lambda_i p\rangle_{q_2}. \tag{A1}$$

Take a homodyne detection on qumode q_2 and if the result is $\lambda_i p = q_s$, then the state collapses to

$$\sum_i \frac{1}{\lambda_i} |\psi_i\rangle \left|\frac{q_s}{\lambda_i}\right\rangle_{p_1}, \tag{A2}$$

where q_2 mode has been discarded. Further, postselecting of $q_s = 0$, then this can realize $g(\lambda) = \frac{1}{\lambda}$. General $g(\lambda)$ in principle can be implemented as follows. First, quantum float computation [6] allows a mapping from $|\lambda_i p\rangle_{q_2}$ to $|1/g(\lambda_i p)\rangle_{q_2}$. Then a homodyne measurement on the qumode q_2 leads to the state

$$\sum_i g(\lambda_i)|\psi_i\rangle \left|\frac{g^{-1}(q_s^{-1})}{\lambda_i}\right\rangle_{p_1}, \tag{A3}$$

which multiplies the coefficient by $g(\lambda_i)$, respectively. It is required that $g^{-1}(q_s^{-1}) = 0$ for some given q_s to disentangle the p_1 qumode. For instance, when $g(x) = x^{-n}$ $(n > 0)$, we can chose $q_s = 0$.

APPENDIX B: FINITE SQUEEZING ANALYSIS

Let us consider squeezed states with squeezing factor s, following the method in Ref. [9]. Then the unitary operator $U = e^{i\eta A^T A \hat{p}_1 \hat{p}_2}$ performs on the prepared state

$$|\Psi\rangle \propto |\psi_A\rangle \int dp_1\, dp_2 s^{-1} e^{-(p_1^2+p_2^2)/2s^2} |p_1\rangle_{p_1} |p_2\rangle_{p_2}. \tag{B1}$$

We emphasize that the factor s^{-1} should not be dropped when analyzing the runtime behavior. It follows by the operator $e^{i\eta\chi\hat{p}_1\hat{p}_2}$ realizing a regularization. To unentangle qumodes from qubits, homodyne detections are conducted on both qumodes with results $q_1 = Q_1$ and $q_2 = Q_2$. Then the quantum state turns to be

$$\sum_i \lambda_i B_i(Q_1, Q_2) \left|\psi_{\mathbf{u}_i}\right\rangle \left|\psi_{\mathbf{v}_i}\right\rangle |Q_1\rangle_{q_1} |Q_2\rangle_{q_2}, \tag{B2}$$

where

$$\begin{aligned} &B_i(Q_1, Q_2) \\ &= \int dp_1\, dp_2 e^{-(p_1^2+p_2^2)/2s^2} e^{i(\alpha_i p_1 p_2 - p_1 Q_1 - p_2 Q_2)} \\ &\propto \frac{\exp\left\{-\left[s^2\left(Q_1^2+Q_2^2\right)+2is^4\alpha_i Q_1 Q_2\right]/2\left(1+s^4\alpha_i^2\right)\right\}}{s\alpha_i\sqrt{1+1/s^4\alpha_i^2}}. \end{aligned} \tag{B3}$$

For brevity we have introduced $\alpha_i = \eta(\lambda_i^2 + \chi)$. Set $\alpha_i^2 s^4 \sim \varepsilon_q^{-1}$, then

$$B_i(Q_1, Q_2) \sim \frac{e^{-(Q_1^2+Q_2^2)/2\alpha_i^2 s^2}}{s\alpha_i}. \tag{B4}$$

The width of distribution of both Q_1 and Q_2 is $\alpha_i s \sim \frac{1}{s\sqrt{\varepsilon_q}}$. For our task it is natural to let the precision ε_q be specified to α_i. We let $\frac{Q_1^2+Q_2^2}{\alpha_i} \lesssim \varepsilon_q$. Note the probability density is $|\lambda_i B_i(Q_1, Q_2)|^2$. Then the success rate of homodyne detection of two qumodes around the center with area ε_q can be approximately calculated through

$$\begin{aligned} &\sum_i \int_{Q_1^2+Q_2^2 \le \alpha_i \varepsilon_q} dQ_1 dQ_2 |\lambda_i B_i(Q_1, Q_2)|^2 \\ &= \sum_i \lambda_i^2 \int_0^{\sqrt{\varepsilon_q \alpha_i}} \frac{e^{-r^2/\alpha_i^2 s^2}}{s^2\alpha_i^2} r\, dr \propto \varepsilon_q^{\frac{3}{2}}. \end{aligned} \tag{B5}$$

The runtime is $O(\varepsilon_q^{-\frac{3}{2}})$, and further if amplitude amplification [42] can be applied to continuous variables, then the runtime would reduce to $O(\varepsilon_q^{-\frac{3}{4}})$.

APPENDIX C: PHYSICAL IMPLEMENTATION: A DEMONSTRATION

For the purpose of demonstration, we consider a data set that contains four training samples, each having two

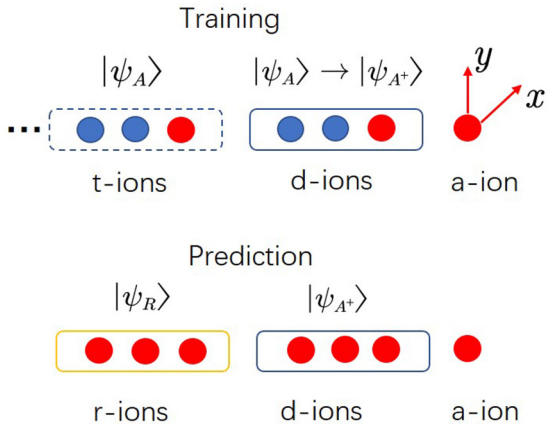

FIG. 3. Ions (in red color) involved at the training and prediction steps.

features. This corresponds to $M = 4$ and $N = 2$. Features for each sample are represented by $\mathbf{a}^{(m)} = (a_0^{(m)}, a_1^{(m)})$ $(m = 0, 1, 2, 3)$, and the corresponding target value is $y^{(m)}$. The regression task is to predict the target value for new data $\tilde{\mathbf{a}} = (\tilde{a}_0, \tilde{a}_1)$.

We choose trapped ions [30–35] for a physical implementation of the algorithm. Our setup contains ions in a Pauli trap lined in the z direction. Each ion can provide a qubit and two qumodes (motion modes in x and y directions). According to their roles in the algorithm, we can group them as follows:

(1) d-ions contain three ions for loading the data information into a quantum state $|\psi_A\rangle = \sum_{m=0}^{3}\sum_{n=0}^{1} a_n^{(m)}|m\rangle|n\rangle$, which would be transformed into the target state $|\psi_{A^+}\rangle$.

(2) t-ions contain three ions. They are initialized to $|\psi_A\rangle$ and help to construct U for the phase estimation. Many copies of t-ions may be required depending on the desired precision. Here we simply choose two copies for the purpose of demonstration.

(3) a-ion provides an ancillary qubit and two qumodes.

(4) r-ions contain three ions for encoding the reference state $|\Psi_R\rangle$.

In total 13 ions are required. We briefly discuss how those ions play their role in the procedures of the algorithm, as illustrated in Fig. 3. First, the a-ion, d-ions, t-ions, and r-ions are initialized to $|0\rangle_a|G_{12}\rangle$, $|\psi_A\rangle$, $|\psi_A\rangle$, and $|\psi_R\rangle$, respectively. In the training step, which transfers $|\psi_A\rangle$ to $|\psi_{A^+}\rangle$, the t-ions, d-ions, and a-ion are involved to perform the quantum phase estimation. For either t-ions or d-ions, only the qubit recording features are involved, as marked in red in Fig. 3. Then a controlled-phase gate and homodyne measurements perform on two qumodes of the a-ion, realizing regularization and singular value transformation, respectively. In the prediction, a swap test involves r-ions (encoding $|\Psi_R\rangle$), d-ions (encoding $|\Psi_A^+\rangle$), and a-ion (providing an ancillary qubit). Note that two motion modes of the a-ion are not required in this stage anymore.

We work on squeezing states of momentum, as a regularization operator $e^{i\eta\chi\hat{x}\hat{y}}$ is more realizable in a single ion.

The implementation details with the five procedures are given as follows:

(1) *State preparation*. d-ions are prepared in the state by $D_A|0\rangle_{d_1}|0\rangle_{d_2}|0\rangle_{d_3} = |\psi_A\rangle$. An explicit quantum circuit

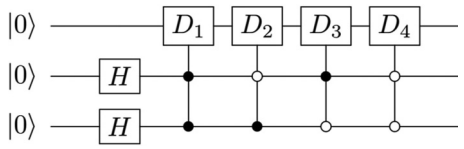

FIG. 4. Quantum circuit for preparing $|\psi_A\rangle$. Here $D_m = e^{i\theta_m\sigma_y}$ $(m = 0, 1, 2, 3)$ with $\theta = \arctan(a_1^{(m)}/a_0^{(m)})$ are single-qubit gates, which realize $D_m|0\rangle = a_0^{(m)}|0\rangle + a_1^{(m)}|1\rangle$. H is the Hadamard gate. Two controls are added to D_m, such that they generate $|m\rangle(a_1^{(m)}|0\rangle + a^{(m)}|1\rangle)$, respectively. We have defined $|0\rangle \equiv |00\rangle$, $|1\rangle \equiv |10\rangle$, $|2\rangle \equiv |01\rangle$, $|3\rangle \equiv |11\rangle$. The quantum circuit outputs the required state $|\psi_A\rangle = \sum_{m=0}^{3}\sum_{n=0}^{1} a_n^{(m)}|m\rangle|n\rangle$.

of D_A can be seen in Fig. 4. t-ions are also prepared in state $|\psi_A\rangle$. For the a-ion, the qubit state is initialed as $|0\rangle_a$, and the motion state is initialized as $|G_{12}\rangle = |0, s\rangle_{p_x}|0, s\rangle_{p_y}$, where the squeezing state $|0, s\rangle_{p_\alpha}$ is obtained by performing a squeezing operator $S_\alpha(1/s) = e^{\frac{i}{2}\ln s(\hat{q}_\alpha\hat{p}_\alpha + \hat{p}_\alpha\hat{q}_\alpha)}$ on the vacuum state $|0\rangle_{c_\alpha}$. r-ions are initialized in the reference state $|\psi_R\rangle = D_{\mathbf{y}} \otimes D_{\tilde{\mathbf{a}}}|0_1 0_2\rangle_r|0\rangle_{r_3} = |\psi_{\mathbf{y}}\rangle|\psi_{\tilde{\mathbf{a}}}\rangle$. Here $|0_1 0_2\rangle_r$ short-notes $|0\rangle_{r_1}|0\rangle_{r_2}$. $D_{\mathbf{y}}$ is a two-qubit operator that prepares $|\psi_{\mathbf{y}}\rangle = y^{(0)}|0_1 0_2\rangle_r + y^{(1)}|1_1 0_2\rangle_r + y^{(2)}|0_1 1_2\rangle_r + y^{(3)}|1_1 1_2\rangle_r$, and $D_{\tilde{\mathbf{a}}}|0\rangle_{r_3} = \tilde{a}_0|0\rangle_{r_3} + \tilde{a}_1|1\rangle_{r_3}$.

(2) *Quantum phase estimation*. This step applies an unitary operation $U = e^{i\eta A^T A\hat{q}_x\hat{q}_y}$, using $\mathrm{Tr}_\rho(e^{i\delta t\eta\hat{q}_x\hat{q}_y S_m}\rho \otimes \rho' e^{-i\delta t\eta\hat{q}_x\hat{q}_y S_m})$. Here $\rho \equiv A^T A = \mathrm{Tr}_{t_1 t_2}|\psi_A\rangle\langle\psi_A|$ is a state on t-ions, and $\rho' \equiv |\psi_A\rangle\langle\psi_A|$ is a state on d-ions. The conditional exponential swap operator $e^{i\delta t\eta\hat{q}_x\hat{q}_y S}$ is $C_S H_a R_a^z(\delta t\eta\hat{q}_x\hat{q}_y)H_a C_S$. Here C_S is a conditional swap operator performing on the t_3 qubit and d_3 qubit, conditioned on the qubit of the a-ion. We emphasize that recently the conditional swap operator has been realized experimentally in trapped ions [36]. $R_a^z(\delta t\eta\hat{q}_x\hat{q}_y) = e^{i\delta t\eta\sigma_z\hat{q}_x\hat{q}_y}$ performs on the a-ion. Since $H_1 = g\hat{q}_x\sigma_x$ and $H_2 = g\hat{q}_y\sigma_y$ can be realized using lasers when simulating Dirac equations [34,35], then a quantum evolution $e^{iH_2\delta t}e^{iH_1\delta t}e^{-iH_2\delta t}e^{-iH_1\delta t}$ can realize $R_h^z(\delta t\eta\hat{q}_x\hat{q}_y)$ by repeating this evolution $\eta/(g^2\delta t)$ times.

(3) *Regularization*. The regularization operator is a controlled-phase gate $C_z = e^{i\chi\eta\hat{q}_x\hat{q}_y}$. This may be obtained using the relation $e^{ik_0 t(\hat{q}_x+\hat{q}_y)^2}e^{-ik_0 t\hat{q}_x^2}e^{-ik_0 t\hat{q}_y^2} = e^{i2k_0 t\hat{q}_x\hat{q}_y}$ by setting $2k_0 t = \chi\eta$. Physically, it can be realized by disturbing the trap for the a-ion as follows. First, relax the trap along both the x and y directions by k_0 in a short time period t, then enhance the trap along the $x = -y$ direction by k_0 in the same period t. If t is small enough, this evolution can approximate C_z once $k_0 = \chi\eta/t$.

(4) *Singular-value transformation*. We need a measurement that projects the momentum of the a-ion in an small area ε_q around the zero point. A modification of the motion tomography [37] can be applied for this task. The projective measurement is $\Pi_a(s) = S^+(1/s)\Pi_a(0)S(1/s)$, where $\Pi_a(0)$ projects to the ion's motional vacuum state $|0\rangle_c$. To achieve this, we first apply $S(1/s)$ in both the x and y directions, then use the quantum jump method to get the probability of a successful projection $\Pi_a(s)$. A projection onto a small area around zero momentum may be approximated by choosing a slightly bigger squeezing factor, which raises the success rate.

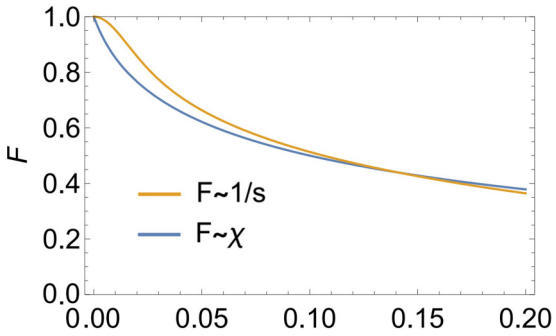

FIG. 5. The relation of fidelity with squeezing $\frac{1}{s}$ (in brown) and with the regularization χ (in blue).

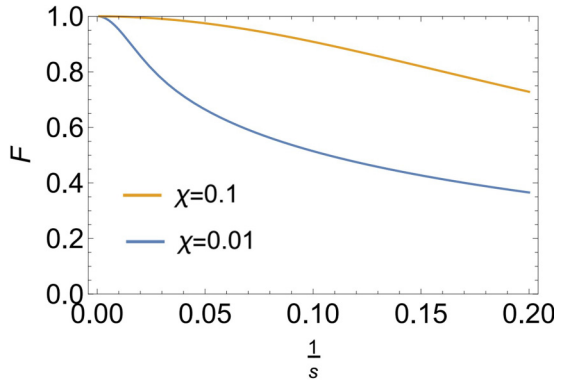

FIG. 6. Fidelity with squeezing under different regularization $\chi = 0.01, 0.1$.

(5) *Prediction*. The swap test requires a conditional swap operation, which swaps states between d-ions and r-ions, conditioned on the qubit in the a-ion. The qubit in the a-ion should be reset in the $|+\rangle_a$ state. After the swap operation, a Hadamard gate performs on the qubit of the a-ion and projects it onto the $|0\rangle_a$ state. The probability of successful projection is $p = \frac{1}{2}(1 + |\langle\psi_{A^+}|\psi_R\rangle|^2)$, and we have $|\tilde{y}| = \sqrt{2p-1}$. The sign of $\tilde{y}$ should be determined by other means.

APPENDIX D: ERROR DUE TO FINITE SQUEEZING AND REGULARIZATION

We analyze the effect of finite squeezing and regularization, by calculating the fidelity between the $|\psi_{A_0^+}\rangle = \sum_i \frac{1}{\lambda_i}|\psi_{\mathbf{u}_i}\rangle|\psi_{\mathbf{v}_i}\rangle$ and $|\psi_{A'}\rangle$. The fidelity turns to be $F(s,\chi) \sim \sum_i f_i(Q_1, Q_2)\frac{1}{\alpha_i}$ (not normalized). Without loss of generality, we assume λ_i^2 distributes uniformly in the zone $[\delta_0, 1]$, where δ_0 is a small positive number to avoid divergence brought by $1/\lambda$. Then the fidelity can be approximated as

$$F(s,\chi) \propto \int_{\delta_0}^{1} d\lambda \exp\left[-\frac{\epsilon_q}{2\eta^2(\lambda+\chi)^2 s^2}\right]\frac{1}{\eta(\lambda+\chi)} = E_i\left[-\frac{\epsilon_q}{2(\delta_0+\chi)^2 s^2\eta^2}\right] - E_i\left[-\frac{\epsilon_q}{2(1+\chi)^2 s^2\eta^2}\right], \tag{D1}$$

where $E_i(z) = -\int_{-z}^{\infty} dt e^{-t^2}/t$. The fidelity increases when s increases as one should expect, and decreases when χ increases, as seen in Fig. 5.

We also fixed the regularization and study the behavior of fidelity with $\frac{1}{s}$ between $|\psi_A^+\rangle$ and $|\psi_A'\rangle$ under different χ. As seen in Fig. 6, with larger regularization the fidelity can decrease slower when reducing the squeezing factor s.

[1] A. Furusawa and P. Van Loock, *Quantum Teleportation and Entanglement: A Hybrid Approach to Optical Quantum Information Processing* (John Wiley & Sons, New York, 2011).
[2] U. L. Andersen, J. S. Neergaard-Nielsen, P. van Loock, and A. Furusawa, Nat. Phys. **11**, 713 (2015).
[3] S. Lloyd, arXiv:quant-ph/0008057.
[4] P. van Loock, W. J. Munro, K. Nemoto, T. P. Spiller, T. D. Ladd, S. L. Braunstein, and G. J. Milburn, Phys. Rev. A **78**, 022303 (2008).
[5] T. Proctor, M. Giulian, N. Korolkova, E. Andersson, and V. Kendon, Phys. Rev. A **95**, 052317 (2017).
[6] S. Lloyd and S. L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999).
[7] N. N. Liu, J. Thompson, C. Weedbrook, S. Lloyd, V. Vedral, M. L. Gu, and K. Modi, Phys. Rev. A **93**, 052304 (2016).
[8] H. K. Lau and M. B. Plenio, Phys. Rev. Lett. **117**, 100501 (2016).
[9] H. K. Lau, R. Pooser, G. Siopsis, and C. Weedbrook, Phys. Rev. Lett. **118**, 080501 (2017).
[10] S. Aaronson, Nat. Phys. **11**, 291 (2015).
[11] C. M. Bishop, *Pattern Recognition and Machine Learning*, Vol. 1 (Springer, Cambridge, 2006).
[12] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature (London) **549**, 195 (2017).
[13] N. Wiebe, D. Braun, and S. Lloyd, Phys. Rev. Lett. **109**, 050505 (2012).
[14] M. Schuld, I. Sinayskiy, and F. Petruccione, Contemp. Phys. **56**, 172 (2015).
[15] S. Lloyd, M. Mohseni, and P. Rebentrost, Nat. Phys. **10**, 631 (2014).
[16] P. Rebentrost, M. Mohseni, and S. Lloyd, Phys. Rev. Lett. **113**, 130503 (2014).
[17] V. Dunjko, J. M. Taylor, and H. J. Briegel, Phys. Rev. Lett. **117**, 130501 (2016).
[18] S. Lloyd, S. Garnerone, and P. Zanardi, Nat. Commun. **7**, 10138 (2016).
[19] G. M. Wang, Phys. Rev. A **96**, 012335 (2017).
[20] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[21] B. D. Clader, B. C. Jacobs, and C. R. Sprouse, Phys. Rev. Lett. **110**, 250504 (2013).
[22] J. Pan, Y. D. Cao, X. W. Yao, Z. K. Li, C. Y. Ju, H. W. Chen, X. H. Peng, S. Kais, and J. F. Du, Phys. Rev. A **89**, 022313 (2014).

[23] M. Schuld, I. Sinayskiy, and F. Petruccione, Phys. Rev. A **94**, 022342 (2016).
[24] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **78**, 052310 (2008).
[25] F. Y. Hong, Y. Xiang, Z. Y. Zhu, L. Z. Jiang, and L. N. Wu, Phys. Rev. A **86**, 010306 (2012).
[26] K. R. Patton and U. R. Fischer, Phys. Rev. Lett. **111**, 240504 (2013).
[27] 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, Phys. Rev. Lett. **114**, 110504 (2015).
[28] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).
[29] S. Kimmel, C. Y.-Y. Lin, G. H. Low, M. Ozols, and T. J. Yoder, npj Quantum Inf. **3**, 13 (2017).
[30] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
[31] H.-K. Lau and D. F. V. James, Phys. Rev. A **85**, 062329 (2012).
[32] C. Shen, Z. Zhang, and L. M. Duan, Phys. Rev. Lett. **112**, 050504 (2014).
[33] L. Ortiz-Gutiérrez, B. Gabrielly, L. F. Muñoz, K. T. Pereira, J. G. Filgueiras, and A. S. Villar, Opt. Commun. **397**, 166 (2017).
[34] L. Lamata, J. Leon, T. Schatz, and E. Solano, Phys. Rev. Lett. **98**, 253005 (2007).
[35] R. Gerritsma, G. Kirchmair, F. Zahringer, E. Solano, R. Blatt, and C. F. Roos, Nature (London) **463**, 68 (2010).
[36] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, Phys. Rev. A **98**, 052334 (2018).
[37] J. F. Poyatos, R. Walser, J. I. Cirac, P. Zoller, and R. Blatt, Phys. Rev. A **53**, R1966 (1996).
[38] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).
[39] D. Kienzler, H. Y. Lo, B. Keitch, L. de Clercq, F. Leupold, F. Lindenfelser, M. Marinelli, V. Negnevitsky, and J. P. Home, Science **347**, 53 (2015).
[40] S. C. Burd, R. Srinivas, J. J. Bollinger, A. C. Wilson, D. J. Wineland, D. Leibfried, D. H. Slichter, and D. T. C. Allcock, arXiv:1812.01812.
[41] P. C. Hansen, BIT Numer. Math. **27**, 534 (1987).
[42] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, arXiv:quant-ph/0005055.