

# Realizing quantum linear regression with auxiliary qumodes

Dan-Bo Zhang,<sup>1</sup> Zheng-Yuan Xue,<sup>1</sup> Shi-Liang Zhu,<sup>1,2,\*</sup> and Z. D. Wang<sup>3,†</sup>

<sup>1</sup>*Guangdong Provincial Key Laboratory of Quantum Engineering and Quantum Materials, and School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China*

<sup>2</sup>*National Laboratory of Solid State Microstructures,*

*School of Physics, Nanjing University, Nanjing 210093, China*

<sup>3</sup>*Department of Physics and Center of Theoretical and Computational Physics, The University of Hong Kong, Pokfulam Road, Hong Kong, China*

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 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.

## 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 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 that 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 consists 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 more rare 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 post-selection. 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 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.

## 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 quantum algorithm. Let us first intro-

\* slzhu@nju.edu.cn

† zwang@hkucc.hku.hk

duce the linear regression. Given a training dataset 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-squares errors,

$$\mathbf{w} = \min_{\mathbf{w}} \sum_{m=1}^M (\mathbf{w}^T \mathbf{a}^{(m)} - y^{(m)})^2 + \chi \|\mathbf{w}\|^2. \quad (1)$$

Here  $\chi \|\mathbf{w}\|^2$  is a regularization term with parameter  $\chi$ , 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  $\chi$  term) reads  $A^+ = (A^T A + \chi I)^{-1} A^T$ . Here  $I$  is the identity matrix. It is inspiring to study the linear regression by using the singular value decomposition of  $A$  by [23],  $A = \sum_i \lambda_i \mathbf{u}_i \mathbf{v}_i^T$ . Here  $\lambda_i$  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  $\tilde{y} = \sum_i \frac{\lambda_i}{\lambda_i^2 + \chi} \mathbf{u}_i^T \tilde{\mathbf{a}} \mathbf{v}_i$ .

To formulate a quantum version of linear regression, we need to encode each component as a quantum state or a quantum operator. For vector  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ , 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  $|n\rangle = |n_1 n_2 n_3 \dots\rangle$ , where  $n = n_1 n_2 n_3 \dots$  is a binary representation of integer  $n$ , and it requires a number of  $\lceil \log_2 N \rceil$  qubits. Accordingly,  $\mathbf{y}$  and  $\tilde{\mathbf{a}}$  are encoded as  $|\psi_{\mathbf{y}}\rangle = \sum_m y^{(m)} |m\rangle$  and  $|\psi_{\tilde{\mathbf{a}}}\rangle = \sum_n \tilde{a}_n |n\rangle$ , 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 an 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]. Firstly,  $A^T$  is not a square matrix, and the Hilbert space should be extended to define a square matrix that contains  $A^T$ . Secondly, two HHL-like algorithms for applying  $A^T$  and  $(A^T A + \chi I)^{-1}$  sequentially are required. Thirdly, both  $A^T$  and  $(A^T A + \chi I)^{-1}$  are not unitary in general, there must be some non-unitary 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. Accord-

ingly, we may write the quantum state 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 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 of  $g(\lambda) = \frac{\lambda}{\lambda^2 + \chi}$ . 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 stats of parameters  $|\psi_{\mathbf{w}}\rangle$  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  $|\psi_{\tilde{\mathbf{a}}}\rangle |\psi_{\mathbf{y}_t}\rangle$  ( $t = 1, 2$  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

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

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 for lots of 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 all-qubit approach. Firstly, quantum phase estimation is

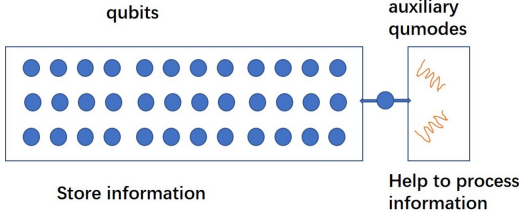


FIG. 1. Illustration of the hybrid approach that exploits the best of both qubits and qumodes. In our implementation of quantum linear regression qubits and qumodes are minimal coupled by connecting only to an ancillary qubit.

applied that registers  $\lambda_i$  in ancillary qubits; secondly, coefficients  $\frac{1}{\lambda_i}$  are obtained by a conditional rotation on an extra ancillary qubit and then a projection to  $|1\rangle$  state. The success rate of projecting to  $|1\rangle$  is proportional to the conditional number  $\kappa = \frac{\lambda_{\max}^2}{\lambda_{\min}^2}$ , where  $\lambda_{\max/\min}$  stands for largest/smallest singular values of  $A$ . Under bad conditional number, e.g.,  $\kappa_c \sim O(N)$  (which naturally arise for low-rank  $A^T A$ ), 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 resource, especially on the number of qubits. For the quantum phase estimation, firstly, series of  $\{e^{iA^T A \tau}\}$  with different time  $\tau$  should be constructed. Secondly, since  $\lambda_i$  are typically continuous numbers, to encode them with precision  $\varepsilon$  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 lots of 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 for alternative approach that requires less resource 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, as would 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 can not be infinite physically. Nevertheless, we first present the algorithm for the case of infinite squeezing, as the main procedures of the algorithm is 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  $|\psi_{\mathbf{A}}\rangle = \sum_m |m\rangle |\psi_{\mathbf{a}^{(m)}}\rangle$  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  $U = e^{i\eta A^T A \hat{p}_1 \hat{p}_2}$  on  $|\psi_{\mathbf{A}}\rangle |0\rangle_{q_1} |0\rangle_{q_2}$  which leads to

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

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

3. *Regularization.* Perform  $e^{i\eta \chi \hat{p}_1 \hat{p}_2}$  that shifts  $q_2$  mode to  $|\eta(\lambda_i^2 + \chi)p\rangle_{q_2}$ .
4. *Singular-value transformation.* After a homodyne detection on qumode  $q_2$  with result  $q_s = 0$ , the state collapses to  $|\psi_{\mathbf{A}^+}\rangle$ , 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  $\tilde{y}'$ , which can be implemented using the method [16, 27]. Introduce an ancillary qubit and construct an entangled state as following  $|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle |\Psi_R\rangle + |1\rangle |\psi_{\mathbf{A}^+}\rangle)$ . Then, a  $\sigma_x$  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 implementable, and success rate for post-selecting  $q_s = 0$  is vanishing.

We now take finite squeezing into account, which leads to a modification of the algorithm in three places. *Firstly*, 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  $|q\rangle = \frac{1}{\sqrt{2\pi}} \int dp e^{-iqp} |p\rangle$ ,  $|G_{12}\rangle = |0\rangle_{q_1} |0\rangle_{q_2}$  at  $s \rightarrow \infty$  limit. *Secondly*, in the singular-value transformation, to fully disentangle qubits and qumodes, two qumodes are post-selected as  $q_1 = Q_1$  and  $q_2 = Q_2$  using homodyne detection. To get a nonzero success rate, we can let the point  $(Q_1, Q_2)$  locates within a small area proportional to  $\epsilon_q$  centered at  $(0, 0)$  (see Appendix B). Now, the unnormalized state can be written as

$$|\psi_{A'}\rangle = \sum_i \lambda_i B_i(Q_1, Q_2) |\psi_{\mathbf{u}_i}\rangle |\psi_{\mathbf{v}_i}\rangle \quad (3)$$

where  $B_i(Q_1, Q_2) \sim e^{-(Q_1^2 + Q_2^2)/2\alpha_i^2 s^2} / s\alpha_i$  at the limit  $\alpha_i^2 s^4 \sim \epsilon_q^{-1}$ . Here  $\alpha_i = \eta(\lambda_i^2 + \chi)$ . Compared with  $|\psi_{A+}\rangle$ , each coefficient now is multiplied by  $B_i(Q_1, Q_2)$  correspondingly. When  $\epsilon_q \rightarrow 0$ , the state  $|\psi_{A'}\rangle$  reduces to  $|\psi_{A+}\rangle$ . In this sense  $\epsilon_q$  is an error. We may take  $B_i(Q_1, Q_2)$  as an extra regularization due to finite squeezing, which will be discussed in Sec. V. Thirdly, 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 with 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 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 of  $U$  involves a density matrix exponentiation of  $A^T A$  using the method in Ref. [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 exponentiation  $e^{i\rho t}$  of density  $\rho = A^T A$  can be obtained by repeatedly applies the exponential swap operation and tracking out  $\rho$ ,

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

Here  $S_m$  is the swap operator. To construct  $U$ , the exponential swap operator should be modified as

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

Here  $R_a^x(\delta t \eta \hat{p}_1 \hat{p}_2) = e^{i\delta t \eta \hat{q}_x \hat{q}_y \sigma_x}$ , and  $C_S = C_S^{bb'} C_S^{dd'} \dots$  is a control swap operator on multi-modes, where  $C_S^{\alpha\alpha'}$  swap between  $\alpha$  and  $\alpha'$ , conditioned on the ancillary qubit  $a$ .

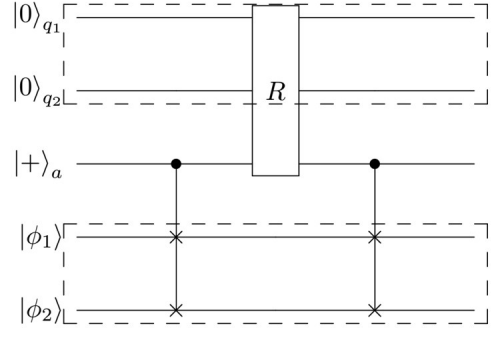


FIG. 2. Quantum circuit for implementing modified exponential swap gate. An ancillary qubit is prepared at  $|+\rangle$ . Qumodes and qubits are not directly coupled.

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 Control-NOT gate that takes  $b(b')$  as the control(target), and  $T_{ab'b}$  is the Toffoli gate that 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

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

Set  $\rho = A^T A$  we obtain  $e^{i\delta t \eta A^T A \hat{p}_1 \hat{p}_2}$ . To construct  $U$  with desired precision  $\epsilon$  it requires  $O(\delta t^{-2} \epsilon^{-1})$  copies of  $\rho$ , which turns to be  $O(\epsilon^{-3})$  for nonsparse and low-rank  $A^T A$  [15].

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], using 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  $H_2 = g \hat{p}_2 \sigma_x$ . Repeat the evolution  $\frac{1}{g^2 \delta t}$  times then we have  $R_a(\delta t \eta \hat{p}_1 \hat{p}_2)$ .

We briefly discuss how to implement the algorithm physically in trapped ions. It is emphasized that the hybrid approach is more feasible for a physical implementation, comparing to the all-qubit approach. For trapped ions, the internal states and motions of trapped ions can respectively serve as discrete and continuous variables for encoding information, and the shape of the trap, external lasers as well as the interactions between ions can be utilized or controlled to manipulate them [30–35]. The setup contains ions in a Pauli trap lined in the  $z$  direction. A specified ion, called 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  $\psi_R$  are prepared using internal states of ions. In the phase estimation, the conditional exponential swap operator turns to be  $e^{i\delta t \eta \hat{q}_x \hat{q}_y S_m} = C_S H_a R_a^z(\delta t \eta \hat{q}_x \hat{q}_y) H_a C_S$ , 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  $H_1 = g\hat{q}_x\sigma_x, H_2 = g\hat{q}_y\sigma_y$ , 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  $x = -y$  direction for 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 approximate 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  $|1\rangle_a$  state. The sign of  $\tilde{y}$  should be determined by other means [23].

## V. DISCUSSIONS 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. Firstly, 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). Secondly, 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 would be revealed below.

We first analyze the runtime of the hybrid approach using two qumodes. It mainly arises from the construction of operator  $U = e^{i\eta A^T A p_1 \hat{p}_2}$  for the quantum phase estimation, and the success rate of homodyne detection at the step of singular value transformation. To construct  $U$ , it requires  $O(t^2 \varepsilon)$  copies of density matrix  $A^T A$  with desired precision  $\varepsilon$ . Meanwhile, each copy of density matrix  $A^T A$  can be obtained from  $|\psi_A\rangle$  that takes runtime  $O(\log(MN))$  using qRAM. For non-sparse and low-rank  $A^T A$ , estimating singular value to precision  $\varepsilon$  by quantum phase estimation takes time  $t = O(\varepsilon^{-1})$  [15]. Thus the runtime in the phase estimation is  $O(\varepsilon^{-3} \log(MN))$ . The runtime for singular value transformation turns to be  $O(\varepsilon_q^{-\frac{3}{2}})$  (see Appendix B). In total, the runtime scales as  $O(\varepsilon_q^{-\frac{3}{2}} \varepsilon^{-3} \log(MN))$ .

We investigate the role of the regularization parameter  $\chi$  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 bad conditional 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  $\epsilon_q$  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 resource of squeezing.

We also mention the shrinkage effects [39] due to the regularization and finite squeezing. Firstly, 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  $B_i(Q_1, Q_2)$ , correspondingly. This further reduces the weighting of low variance components. Rather than taking  $B_i(Q_1, Q_2)$  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 would 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 post-selection. The smaller  $s$ , the bigger  $\epsilon_q$ , and thus the higher the success rate. On the other hand, smaller  $s$  leads to a bigger shrinkage. Thus, a proper squeezing factor  $s$  is not only necessary for efficiently implementing the quantum algorithm, but also may be beneficial for machine learning.

In summary, we adopt 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 approaches have the same order of runtime as the all-qubit regarding to the dimension of data  $N$ , the number of training data (samples)  $M$ , and the precision  $\varepsilon$ , but can save the using of ancillary qubits at the price of requiring finite squeezing states with qumodes. We also demonstrates the important role of finite squeezing for efficiently running the algorithm. Moreover, we show that the shrinkage effect due to finite squeezing may provide a new type of regularization for linear regression. We wish 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 NKRD of China (Grant No. 2016YFA0301800), the NSFC (Grants No. 91636218, No.11474153).

- 
- [1] A. Furusawa and P. Van Loock, *Quantum teleportation and entanglement: a hybrid approach to optical quantum information processing* (John Wiley & Sons, 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 (2000).
- [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, 2006).
- [12] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **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, Nature Communications **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** (2008), ARTN 052310 10.1103/PhysRevA.78.052310.
- [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. Garca-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 Information **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. Zähringer, E. Solano, R. Blatt, and C. F. Roos, Nature **463**, 68 (2010).
- [36] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, arXiv:1712.08581 (2017).
- [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] P. C. Hansen, BIT Numerical Mathematics **27**, 534 (1987).
- [40] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, arXiv:quant-ph/0005055 (2000).

## Appendix A: Qumodes for quantum computing

*Quantum phase estimation.* At the heart of QPE is quantum Fourier transformation which gives exponential speed-up, as performing QFT on  $N$  qubits takes the order of  $N^2$  quantum operations, while fast Fourier transformation takes  $O(N^2)$ . 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, quantum phase estimation [3] routine can be much simplified. Note  $H|e_i\rangle = E_i|e_i\rangle$ . To write eigenvalues into the registering qumode, one just need to performs 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]

$\varepsilon^{-1}$  is required for desired precision  $\varepsilon$ . In contrast, for all-qubit system, registering  $E_i$  with same precision requires  $O(\log \varepsilon^{-1})$  qubits.

*Quantum addition.* Consider  $a + b$ . First encode  $a$  as  $|a\rangle_q$ , then the addition is 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  $\lambda_i$  relates to  $|\psi_i\rangle$ , such as a pair of eigenvalue and eigenstate. Here quantum amplitudes are mapped as  $a_i \rightarrow a_i g(\lambda_i)$ . We show how this can be achieved using entangled qumodes and projection. This is unique for entangled states of continuous-variable. For illustration we give the derivation only for infinite squeezing. We assume that the following quantum state can be prepared,

$$|\psi\rangle = \sum_i |\psi_i\rangle \int dp |p\rangle_{p_1} |\lambda_i p\rangle_{q_2}. \quad (\text{A1})$$

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

$$\sum_i \frac{1}{\lambda_i} |\psi_i\rangle \left| \frac{q_s}{\lambda_i} \right\rangle_{p_1}, \quad (\text{A2})$$

where  $q_2$  mode has been discarded. Further, post-selecting of  $q_s = 0$ , then this can realize  $g(\lambda) = \frac{1}{\lambda}$ . General  $g(\lambda)$  in principle can be implemented as follows. Firstly, 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

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

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

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

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

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

where

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

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

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

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  $\varepsilon_q$  specified to  $\alpha_i$ . We let  $\frac{Q_1^2 + Q_2^2}{\alpha_i^2} \lesssim \varepsilon_q$ . 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  $\varepsilon_q$  can be approximately calculated through

$$\sum_i \int_{Q_1^2 + Q_2^2 \leq \alpha_i \varepsilon_q} dQ_1 dQ_2 |\lambda_i B_i(Q_1, Q_2)|^2 \\ = \sum_i \lambda_i^2 \int_0^{\sqrt{\varepsilon_q \alpha_i}} \frac{e^{-r^2/\alpha_i^2 s^2}}{s^2 \alpha_i^2} r dr \propto \varepsilon_q^{\frac{3}{2}} \quad (\text{B5})$$

The runtime is  $O(\varepsilon_q^{-\frac{3}{2}})$ , and further if amplitude amplification [40] can be applied to continuous variables then the runtime would reduce to  $O(\varepsilon_q^{-\frac{3}{4}})$ .