Maximal Entropy Approach for Quantum State Tomography

Rishabh Gupta,1 Rongxin Xia,1 Raphael D. Levine,2,3 and Sabre Kais4,*

1 Department of Chemistry, Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana, USA
2 The Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
3 Department of Chemistry and Biochemistry and Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA
4 Department of Chemistry, Department of Physics and Astronomy, and Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, Indiana, USA

(Received 31 August 2020; accepted 24 December 2020; published 2 February 2021)

Quantum computation has been growing rapidly in both theory and experiments. In particular, quantum computing devices with a large number of qubits have been developed by IBM, Google, IonQ, and others. The current quantum computing devices are noisy intermediate-scale quantum devices, and so approaches to validate quantum processing on these quantum devices are needed. One of the most common ways of validation for an n-qubit quantum system is quantum tomography, which tries to reconstruct a quantum system’s density matrix by a complete set of observables. However, the inherent noise in the quantum systems and the intrinsic limitations pose a critical challenge to precisely know the actual measurement operators that make quantum tomography impractical in experiments. Here, we propose an alternative approach to quantum tomography, based on the maximal information entropy, that can predict the values of unknown observables based on the available mean measurement data. This can then be used to reconstruct the density matrix with high fidelity even though the results for some observables are missing. Of additional contexts, a practical approach to the inference of the quantum mechanical state using only partial information is also needed.

DOI: 10.1103/PRXQuantum.2.010318

I. INTRODUCTION

Quantum technologies increasingly require measuring the mean value of observables towards determining a detailed state of a nanosystem. The method of characterization of a quantum system [1–5] by measuring the expectation values of a complete set of observables is known as quantum state tomography [3,5,6] and forms an important basis in quantum computation operations [1,7,8]. A simple example is the case when the system can be described by a wave function spanning N basis states: $|\psi\rangle = \sum_{i=1}^{N} a_i |i\rangle$. The $a_i$ are complex-valued amplitudes and $|a_i|^2 = |\langle i | \psi \rangle|^2$ is the probability of being in quantum basis $|i\rangle$. The exact construction of the density matrix would require measurements of $N^2 - 1$ observables. The finite values of the observables generated by the dynamics allow us to specify $N^2 - 1$ independent unitary matrices, which give rise to the SU(N) Lie algebra [9]. SU(2), for example, is a set of three unitary matrices and is what is generated by a qubit. The SU(2) algebra that describes a two-state quantum mechanical system is the quantum generalization of Boolean algebra and can be realized by its three Pauli generators ($\sigma_x$, $\sigma_y$, and $\sigma_z$).

One would expect on general grounds [10] that if one could make repeated measurements it will be possible to characterize the quantum mechanical state with fewer than $N^2 - 1$ observables. Like the classical noiseless coding theorem of Shannon [1,11], the quantum mechanical version by Schumacher [1,12] does not tell us how to construct a code that will do it. This is one motivation for trying to use fewer than $N^2 - 1$ observables. The other is that in a variety of practical contexts we have a limited number of observations or we have measurements of limited fidelity. This is particularly so when the quantum device is noisy, as in so-called noisy intermediate-scale quantum (NISQ) devices [13]. This may lead to an ambiguous description of a quantum state.

In this paper we discuss measured mean values in the sense of separate measurements on a large number of
independent elementary systems. In statistical physics one is often concerned with measurements on an ensemble, a mixture, of systems. In statistical physics one can therefore provide a physical interpretation of the Lagrange multipliers that arise in the process of seeking a constrained maximum of the entropy. This was made clear in the seminal paper of Jaynes [14], who used pressure as an example. Extensive applications to hydrodynamics were discussed early on by Zubarev [15], and a more recent presentation of this work is given in Ref. [16]. Early application to quantum optics, including the dramatic change in the distribution of photons at the lasing threshold, was carried out by Haken [17]. A more recent discussion of this and many additional examples for a variety of physical systems are given in Ref. [18]. An overview of the field can be found in an edited volume [19]. Full statistical determination of quantum observables characterizing a quantum system is in a low-rank state to reduce unnecessary measurements [24,25] or by using the current measurements to decide the next measurements [26]. The first four of these are probabilities of the four basis states and the remaining twelve are the coherences between them. Following the formalism of the maximal information entropy [19] subject to known average values of certain operators \( \hat{f}_k \), the density operator in terms of the Lagrange multipliers \( \lambda_k \) [39] can be obtained as [14,29,40]

\[
\hat{\rho} = \frac{1}{Z(\lambda_1, \ldots, \lambda_k)} \exp \left\{ -\sum_k \lambda_k \hat{f}_k \right\},
\]

where \( Z(\lambda_1, \ldots, \lambda_k) = \text{Tr}(\exp(-\sum_k \lambda_k \hat{f}_k)) \) insures the normalization as \( \text{Tr}(\rho) = 1 \). Following the motivation for using the maximal entropy formalism, as discussed in Sec. 1, the form of the operator in Eq. (2) is to consider a density operator that is compatible with the known mean values of the measurements. The Lagrange multipliers are thus defined to satisfy consistency with the values of the constraints as well as the maximization of the von Neumann entropy. Equation (2) already assumes that not all observables are available, so let us consider the case where we only know the mean values of a probability and a coherence. The result for the density operator of maximal entropy subject to these constraints is

\[
\hat{\rho} = \frac{1}{Z(\lambda_{11}, \lambda_{12})} \exp\left\{ -\lambda_{11} |1\rangle \langle 1| - \lambda_{12} |1\rangle \langle 2| - \lambda_{12}^* |2\rangle \langle 1| \right\}.
\]

To compute the density matrix explicitly, we first diagonalize the Hermitian matrix \( A \), which is the exponent in
Eq. (3),

\[
\Lambda = \begin{bmatrix}
-\lambda_{11} & -\lambda_{12} & 0 & 0 \\
-\lambda_{12}^* & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix} = \sum_{i=1}^{4} \epsilon_i \langle \phi_i | \phi_i \rangle,
\]

where the \( \{ \epsilon_i, |\phi_i\rangle \} \) correspond to the eigenvalues and eigenvectors of \( \Lambda \). The density operator can now be computed in the basis of eigenvectors of \( \Lambda \):

\[
\hat{\rho} = \frac{1}{Z} \left( \exp(\epsilon_1) |\phi_1\rangle \langle \phi_1| + \exp(\epsilon_2) |\phi_2\rangle \langle \phi_2| + \exp(\epsilon_3) |\phi_3\rangle \langle \phi_3| + \exp(\epsilon_4) |\phi_4\rangle \langle \phi_4| \right),
\]

(4)

\[
Z = \text{tr} (\exp(\Lambda)) = \sum_{i=1}^{4} \exp(\epsilon_i).
\]

(5)

A practical form of Eq. (4) is obtained by expanding the projection operators \(|\phi_i\rangle \langle \phi_i|\) in terms of our initial basis (1). The final form of the density operator is then given by

\[
\hat{\rho} = \frac{1}{Z} \sum_{i} \exp(\epsilon_i) |\phi_i\rangle \langle \phi_i| = \frac{1}{Z} \left[ 4|4\rangle \langle 4| + 3|3\rangle \langle 3| + (a + b)|1\rangle \langle 1| + \left( \frac{a}{k_3^*} + \frac{b}{k_4^*} \right) |1\rangle \langle 2| + \left( \frac{a}{|k_3|^2} + \frac{b}{|k_4|^2} \right) |2\rangle \langle 2| \right],
\]

(6)

where \( Z = \sum_{i} \exp(\epsilon_i) \), \( k_3 = -\epsilon_3/\lambda_{12}^* \), \( k_4 = -\epsilon_4/\lambda_{12}^* \),

\[
a = \frac{|k_3|^2}{\sqrt{(k_3^* + 1)(k_3^* + 1)}}, \quad \text{and}
\]

\[
b = \frac{|k_4|^2}{\sqrt{(k_4^* + 1)(k_4^* + 1)}}.
\]

To determine the values of the Lagrange multipliers (\( \lambda_{11}, \lambda_{12} \)), we use information about the measured mean values of the operators \( x_{11} = |1\rangle \langle 1| \) and \( x_{12} = |1\rangle \langle 2| \), and solve the system of transcendental equations (7), (8) for the unknown Lagrange multipliers. Based on the density operator defined by Eq. (6), we have

\[
x_{11} = f(\lambda_{11}, \lambda_{12}) = \frac{1}{Z} (a + b),
\]

(7)

\[
x_{12} = g(\lambda_{11}, \lambda_{12}) = \frac{1}{Z} \left[ \left( \frac{a}{k_3^*} + \frac{b}{k_4^*} \right) \right],
\]

(8)

where \( f \) and \( g \) are functions of the unknown Lagrange multipliers and equate to the coefficients of the corresponding basis operators in the equation of the density operator (6). A detailed description of the method for obtaining the Lagrange multipliers along with its analysis is provided in the Appendices. As a practical matter, it is much more convenient to solve for the mean measurement values as a function of \( \lambda_{11} \) and \( \lambda_{12} \). An illustration of this procedure is given in Fig. 1. Upon reconstruction of the density matrix, the coefficients of the basis operators are functions of the Lagrange multipliers. These functions equate to the mean values of the measurements, as shown in Eqs. (7) and (8). The variables \( (k_3, k_4, a, b) \) defined in Eq. (6) are all dependent on the eigenvalues \( \epsilon_i \) that are expressed in terms of the Lagrange multipliers, as shown in the Appendices, and, therefore, Eqs. (7) and (8) establish a closed functional relationship between the matrix elements of the density operator and the Lagrange multipliers.

In this paper the constraints in the maximal entropy formalism are operators. The resulting density matrix for the state is then an exponential function of a sum of these operators, each one weighted by its Lagrange multiplier. Taking matrix elements of such an operator is, in general, a nontrivial exercise in Lie algebra. Yet it is necessary to compute the trace of such a density operator with the

FIG. 1. Heat map representations of \( x_{11} \) and \( x_{12} \) as a function of \( \lambda_{11} \) and \( \lambda_{12} \) [keeping \( \text{Im}(\lambda_{12}) = 0 \)]. If there are no constraints then all three quantum states should be equally probable, so that \( x_{11} = \frac{1}{3} \), and there should be no finite coherence, so that \( x_{12} = 0 \). This is seen to be the case when the Lagrange multipliers are all zero. From these maps one can read the values of \( \lambda_{11} \) and \( \lambda_{12} \) for each given value of the observables.
constraints. We choose what is to us an effective way: diagonalizing the exponent. We can thereby readily generate a spectral decomposition of the density. This is a key point in our proposition as, following the diagonalization of the constraint matrix $A$ we reconstruct the density matrix in terms of the initial basis operators, we also obtain mean values of unknown operators and, therefore, diagonalization is unavoidable. But this does not imply that the cost of computation will increase upon increasing the size of the system. This is because we perform partial diagonalization rather than full diagonalization, so the cost of computation is constant. We use the analytical expression for the eigenvalues when two mean measurements $(x_{11}, x_{12})$ are known. This analytical expression is independent of the size of the system, so even for an $n$-qubit system, we can still use the analytical expression for the eigenvalues corresponding to a 2-qubit system provided the known mean measurements are $(x_{11}, x_{12})$. Although diagonalization is intrinsic to our approach of reconstructing density matrices, technically, the Lagrange multipliers are not obtained through diagonalization. However, as a consequence of the reconstruction of the density matrix, we also obtain a system of transcendental equations that need to be solved to determine the unknown Lagrange multipliers. A method for obtaining the Lagrange multipliers was proposed in Ref. [41], and shown in Ref. [20], where the quantum observables are the moments of the photon statistics. We implement the iterative Newton method to determine the Lagrange multipliers for the example discussed in the Appendices. The method is quite sensitive to the initial guess for our case and, therefore, we opt for a more efficient inbuilt PYTHON library scipy to solve the transcendental equations to determine the Lagrange multipliers.

There is another significant consideration why our method of diagonalization of the constraint matrix is to be preferred. The constraints that one deals with in a reconstruction problem are Hermitian operators. These operators typically do not commute. It is a rather challenging algebraic task to factor an exponential matrix of noncommuting operators. It can be done, but it is by far not a simple operation. Our method of diagonalization bypasses this issue completely. We demonstrate this through several detailed analytical examples.

III. DENSITY MATRIX RECONSTRUCTION FOR A SINGLE QUBIT

To theoretically demonstrate how close Eq. (2) is to the true density matrix (that contains all the information), we consider a quantum mechanical system corresponding to a single qubit described by the $SU(2)$ algebra for which the density matrix is reconstructed, knowing the expectation values of the operators $|1\rangle\langle1|$ and $|1\rangle\langle2|$. The density operator for such a system is given by Eq. (3). Following the same procedure, we first diagonalize the Hermitian matrix $A$ corresponding to the exponent in Eq. (3) to obtain the eigenvalues $\{\epsilon_i\}$ and eigenvectors $\{|\phi_i\rangle\}$:

$$A = \begin{bmatrix} -\lambda_{11} & -\lambda_{12} \\ -\lambda_{12}^* & 0 \end{bmatrix} = \sum_{i=1}^{2} \epsilon_i |\phi_i\rangle \langle \phi_i|.$$ 

The density operator is then obtained by expanding the projection operators $|\phi_i\rangle \langle \phi_i|$ in terms of the initial basis operators $\{|1\rangle \langle1|, |1\rangle \langle2|, |2\rangle \langle1|, |2\rangle \langle2|\)$:

$$\hat{\rho} = \frac{1}{Z} \sum_i \exp[\epsilon_i] |\phi_i\rangle \langle \phi_i| = \frac{1}{Z} \left[ (a + b)|1\rangle \langle1| + \left( \frac{a}{k_1^2} + \frac{b}{k_2^2} \right) |1\rangle \langle2| + \left( \frac{a}{|k_1|^2} + \frac{b}{|k_2|^2} \right) |2\rangle \langle1| \right] \left( \frac{1}{k_1^2} + \frac{1}{k_2^2} \right). \tag{9}$$

Here $Z = \sum_i \exp[\epsilon_i]$, $k_1 = -\epsilon_1/\lambda_{12}^*$, $k_2 = -\epsilon_2/\lambda_{12}^*$,

$$a = \frac{|k_1|^2}{\sqrt{(k_1^2 + 1)(k_1^2 + 1)}} \exp[\epsilon_1], \quad \text{and} \quad b = \frac{|k_2|^2}{\sqrt{(k_2^2 + 1)(k_2^2 + 1)}} \exp[\epsilon_2].$$

From Eqs. (3) and (9), it is evident that, even if we know only two mean measurements $x_{11}$ and $x_{12}$, we do get information about the unknown mean measurement $x_{22}$. From Eq. (9), $x_{22} = (a/k_1^2 + b/k_2^2)/Z$, which is exactly equal to the true mean measurement value of the operator $|2\rangle\langle2|$. This is because the normalization constraint is inherent to Eq. (3) and so, for a single-qubit system, if the probability of one state is known, the other probability is straightforward and the maximal entropy-based density matrix reconstruction leads to an exact result.

IV. IBM Q TEST

To test the performance of our approach, we propose to reconstruct the density matrix in terms of measurements from numerical experiments conducted on IBM quantum computing chips. Several works have already demonstrated reconstructing the density matrix based on measurements from IBM quantum chips. One such work involved characterization of qubit readouts using quantum detector tomography on IBM quantum computers [42]. IBM Q provides several high-performance simulators to test and optimize any quantum circuit and to compare the results with that of real quantum devices. We can compare our reconstruction with the exact density matrix by the quantum circuits that can be derived numerically using the
measurements from IBM Q. IBM Q has provided several quantum chips with 5 to 16 qubits available to the public. The IBM device used to carry out calculations for the current work is IBM Q 5 Yorktown [43].

V. RESULTS AND DISCUSSION

The current work includes reconstruction of the density matrix using the known mean values of a probability \((x_{11})\) and a coherence \((x_{1K})\) to predict an unknown probability \((x_{KK})\) for \(K = 1, 2, \ldots, N\). The proposed theory of density matrix reconstruction using the maximal entropy formalism is tested and validated for 2-qubit and 3-qubit quantum systems using numerical simulations followed by its implementation in IBM’s Qiskit [44].

A. 2-Qubit quantum systems

1. Numerical simulations to test the reconstruction of the density matrix

The 2-qubit quantum system gives rise to the \(SU(4)\) Lie algebra. In order to reconstruct the density matrix for such a system, given a certain number of known mean measurements, and to verify if such a density matrix is able to reproduce the unknown mean measurement values, we consider a numerical simulation experiment. For such an experiment, we consider two cases: case \(A\) represents a quantum system for which the expectation values of a probability \((x^A_{11})\) and a coherence \((x^A_{1K})\) are known; case \(B\) represents the same 2-qubit quantum system but with the added knowledge of one more expectation value \((x^B_{KK})\). For both cases, the maximal entropy formalism is employed to calculate the density matrix of the 2-qubit system. The density matrix for the case \(B\) serves as the original density matrix and so, if the density matrix reconstructed for case \(A\) is able to predict the value of the unknown probability \((x^B_{KK})\) then it can be shown that the tomography of a quantum system can be carried out using the maximal entropy formalism with the knowledge of fewer mean measurement values. Let us look at the two cases in more detail for the specific \(K = 2\) illustration, which can then be generalized to \(K = 3, 4\) for the 2-qubit quantum systems.

Case \(A\) \((x_{11}, x_{12})\) and case \(B\) \((x_{11}, x_{12}, x^B_{22})\) known. To reconstruct the density matrix for both these cases as per Eq. (6), the unknown Lagrange multipliers need to be determined. For the ease of numerical calculations, the Lagrange multipliers are assumed to be real. The unknown Lagrange multipliers can easily be calculated from Eqs. (7) and (8) using the available mean measurements: \((x_{11}, x_{12})\) for case \(A\) and \((x_{11}, x_{12}, x^B_{22})\) for case \(B\). Upon reconstruction of the density matrices, the values of the predicted probabilities are compared with the true values to substantiate the proposed approach. The linear plot between the true and predicted values of \(x_{22}\) in Fig. 2(a) and the plot of their absolute difference in Fig. 2(b) show how close the predicted values \((x^A_{22})\) are to the true values \((x^B_{22})\) of the mean measurements. Thus, the reconstructed density matrix for a 2-qubit system is successful in accurately predicting \(x_{22}\) using just a probability \((x_{11})\) and a coherence \((x_{12})\).

Just like the above case of the prediction of \(x_{22}\) for 2-qubit quantum systems, the procedure to reconstruct the density matrix for the prediction of \(x_{33}\) or \(x_{44}\) is also similar. Here also, to accurately predict \(x_{KK}\), we assume that two known mean measurements are available: \(x_{11}\) and \(x_{1K}\). A similar numerical simulation experiment is performed to test the prediction of the unknown observable \((x_{KK})\). As can be seen in Figs. 3(b) and 3(d), the absolute difference between the predicted and true mean values of \(x_{KK}\) is of the order of \(1 \times 10^{-9}\). The numerical simulation results confirm that the maximal entropy formalism can successfully reconstruct the density matrix with fewer than \(N^2 - 1\) known mean measurements.

2. Implementation in IBM’s Qiskit

Qiskit is an open-source quantum computing software development framework provided by IBM to run quantum
programs on prototype quantum devices in IBM Q. The four different elements of Qiskit, Terra, Aer, Ignis, and Aqua, provide the tools to build and run quantum circuits, simulate device behavior, study and mitigate errors, and solve real-world problems.

To test the reconstruction of density matrix, we first built various quantum circuits by applying different gate models followed by simulating the circuit with the Aer provider in Qiskit. For the Qiskit software stack, Aer provides a high-performance simulator framework with multiple backends to simulate quantum circuits. For the current 2-qubit system, we used three different gate models comprising of Hadamard, control $X$, control $Z$, and rotation gates: $R_x$, $R_y$, $R_z$. We vary the rotation angle $\theta$ of the rotation gates to obtain the various quantum circuits to test our reconstruction. Upon construction of the circuit, the statevector_simulator backend from Aer is used to obtain the final state upon simulation. This serves as a method of obtaining mean measurements without the consideration of errors. From the state vector, one can obtain the mean values of the measurements as follows. Consider the 2-qubit quantum system obtained upon simulation of the circuit as per the given gate model such that the system can be described by the state vector

$$|\psi\rangle = a_0 |00\rangle + a_1 |01\rangle + a_2 |10\rangle + a_3 |11\rangle,$$

where the $a_i$ are complex amplitudes and $|a_i|^2 = |\langle i | \psi \rangle|^2$ correspond to the probabilities of the measurements

$$x_{11} = a_0^* a_0, \quad x_{22} = a_1^* a_1, \quad x_{33} = a_2^* a_2, \quad x_{44} = a_3^* a_3.$$

Since the complex amplitudes for the 2-qubit states are available from the state vector, we can also calculate the coherences

$$a_i^* a_j = \langle \psi | (i) \langle j | \psi \rangle.$$

For example, $x_{12} = a_0^* a_1$, $x_{13} = a_0^* a_2$, etc.

After obtaining the probabilities and the coherences from the state vector, the mean values of the measurements are calculated using the formalism from maximal entropy theory to compare the results. In Fig. 4 we show the three gate models along with the corresponding plots between the true and predicted mean measurements for $x_{kk}$. The linear plots verify that the reconstructed density matrix correctly predicts the value of $x_{kk}$.

### 3. Trace distance

As a measure of the distance between the state reconstructed using the maximal entropy formalism for an incomplete number of known mean measurements and the original state that contains all the information, the trace distance is calculated. When two mean measurements ($x_{11}$, $x_{12}$) are known, we have shown that an accurate prediction of $x_{22}$ can be obtained. Therefore, for an adequate comparison between the reconstructed state and the true state and also to verify the accurate prediction of $x_{22}$, we also require knowledge of one more probability, which can be either $x_{33}$ or $x_{44}$. The density operator corresponding to maximal entropy for the known mean measurements ($x_{11}$, $x_{12}$, $x_{33}$)
The reconstructed density matrix in its practical form in terms of the initial basis operators is

$$\hat{\rho} = \frac{1}{Z(\lambda_{11}, \lambda_{12}, \lambda_{33})} \exp\{-\lambda_{11} |1\rangle \langle 1| - \lambda_{12} |1\rangle \langle 2| - \lambda_{12}^* |2\rangle \langle 1| - \lambda_{33} |3\rangle \langle 3|\}. \tag{13}$$

The trace distance is calculated for a sample 2-qubit quantum circuit consisting of

$$\hat{\rho} = \frac{1}{Z} \left[ |4\rangle \langle 4| + c |3\rangle \langle 3| + (a + b) |1\rangle \langle 1| + \left(\frac{a}{k_3^2} + \frac{b}{k_4^2}\right) |1\rangle \langle 2| + \left(\frac{a}{k_3^2} + \frac{b}{k_4^2}\right) |2\rangle \langle 1| + \left(\frac{a}{k_3^2} + \frac{b}{k_4^2}\right) |2\rangle \langle 2| \right], \tag{14}$$

where $Z = \sum \exp(\epsilon_i)$, $k_3 = -\epsilon_3/\lambda_{12}^*$, $k_4 = -\epsilon_4/\lambda_{12}^*$,

$$a = \frac{|k_3|^2}{\sqrt{(k_3^2 + 1)(k_3^4 + 1)}},$$

$$b = \frac{|k_4|^2}{\sqrt{(k_4^2 + 1)(k_4^4 + 1)}},$$

and $c = \exp(\epsilon_2)$. The $\epsilon_i$ are the eigenvalues corresponding to the Hermitian matrix obtained from the exponent in Eq. (13). The unknown Lagrange multipliers are determined by solving the system of transcendental equations corresponding to the known mean measurements

$$x_{11} = \frac{1}{Z} (a + b), \quad x_{12} = \frac{1}{Z} \left(\frac{a}{k_3^2} + \frac{b}{k_4^2}\right), \quad x_{33} = \frac{c}{Z}. \tag{15}$$

To compare the reconstructed density matrix in Eq. (14) with the true density matrix, the trace distance is calculated for a sample 2-qubit quantum circuit consisting of...
an $R_y$ gate on each qubit to provide the mean measurement values of the required operators. The rotation angle $\theta$ of the $R_y$ gate is varied to obtain different circuits for comparison. In Fig. 5 we show a plot between the trace distance and the true value of $x_{22}$ arising from the quantum circuit. It can be seen from Fig. 5 that the trace distance is of the order of $1 \times 10^{-9}$, which verifies that the reconstructed density matrix from maximal entropy with incomplete information is very close to the true density matrix.

**B. 3-Qubit quantum systems**

The 3-qubit quantum system gives rise to the SU(8) Lie algebra. The density matrix reconstruction for such a system using the maximal information entropy formalism is similar to that for the 2-qubit system, where the SU(4) group specifies the system. Here also, it is assumed that two mean measurements are known ($x_{11}, x_{12}$) and we try to predict $x_{KK}$ using the reconstructed density matrix. The density operator for the 3-qubit system when ($x_{11}, x_{12}$) are the known mean measurements is

$$\hat{\rho} = \frac{1}{Z} \sum_i \exp(\epsilon_i) |\phi_i\rangle \langle \phi_i|$$

where $Z = \sum_i \exp(\epsilon_i)$, $k_3 = -\epsilon_3/\lambda_{12}$, $k_4 = -\epsilon_4/\lambda_{12}$,

$$a = \frac{|k_3|^2}{\sqrt{(k_3^2 + 1)(k_3^2 + 1)}} \exp(\epsilon_3), \quad \text{and}$$

$$b = \frac{|k_4|^2}{\sqrt{(k_4^2 + 1)(k_4^2 + 1)}} \exp(\epsilon_4).$$

The method adopted to validate the theory for the SU(8) case is the same as that for the SU(4) case. Upon simulation of the quantum circuit using the statevector_simulator backend, we obtain the state vector corresponding to the final state:

$$|\psi\rangle = a_0 |000\rangle + a_1 |001\rangle + a_2 |010\rangle + a_3 |011\rangle + a_4 |100\rangle + a_5 |101\rangle + a_6 |110\rangle + a_7 |111\rangle. \quad (17)$$

**FIG. 6.** A sample circuit diagram with variable $\theta$ in $R_y$ gates for the 3-qubit SU(8) case and the corresponding plots of the true ($x_{KK}$) and predicted ($\bar{x}_{KK}$) mean values for $x_{KK}$. The linear plots show that the maximal entropy formalism-based reconstructed density matrix accurately predicts $x_{KK}$. 
The mean values of the measurements, the probabilities and the coherences, are calculated using Eqs. (11) and (12), respectively, which are used to reconstruct the density matrix for various quantum circuits and predict the unknown probability $x_{KK}$, as discussed in the following section for one example gate model.

1. Testing on some complicated circuits

In Fig. 6 we show an example quantum circuit with a gate model comprising of various quantum gates, such as Hadamard, Toffoli, control $X$, control $Z$, and $R_y$ gates. The rotation angle $\theta$ in the $R_y$ gates on qubit 0 and qubit 1 is varied to obtain the different quantum circuits. Upon construction of the circuits, the statevector_simulator backend is performed to obtain the mean values of measurements and the corresponding values for $x_{KK}$ are predicted using the reconstructed density matrix. It can be seen from Fig. 6 that the results for $x_{KK}$ obtained from the simulations match with those predicted by the reconstructed density matrix.

VI. ERROR ANALYSIS

Three techniques to generate inputs for obtaining the mean values of measurements are considered to verify the reconstructed density matrix for the prediction of $x_{KK}$ for 2-qubit and 3-qubit quantum systems. For the first technique, the statevector_simulator backend of the Aer element in Qiskit is employed, as discussed in Eq. (10). The advantage of this approach is that the coherences can be calculated directly from the complex amplitudes of the basis states. Although, this would not help in the estimation of error as, from any experiment on a real quantum device, the exact state cannot be obtained. However, we could verify that the maximal entropy formalism-based density matrix reconstruction accurately predicts the value of the unknown probability by comparing the expectation values from the reconstructed density matrix with the expectation values from the statevector_simulator backend.

The second technique of generating inputs employs the qasm_simulator backend in the Aer provider in Qiskit, wherein only the probabilities can be obtained directly, and to obtain the coherences, the corresponding operator is decomposed into tensor products of Pauli matrices, as discussed in the Appendices. The calculation of coherences using the decomposition method results in statistical errors that are inversely proportional to the square root of the number of times a circuit is run on the simulator. These statistical errors are reflected in the results of the simulation of the circuit using the qasm_simulator backend.

Finally, the density matrix is reconstructed using the obtained mean measurement values from the IBM machine for the considered quantum circuit using the 5-qubit quantum chip with the ibmqx2 backend [43]. From the IBM machine experiment, we can also only obtain the probabilities and, therefore, again the coherences are obtained using the decomposition method described in the Appendices. Apart from this, we also used Qiskit’s ignis.mitigation.measurement module in order to mitigate the measurement errors. This is carried out by constructing a calibration matrix for the measurements and thereby mitigating the measurement errors.

As can be seen from Fig. 7, for the sample circuit considered, the maximal entropy formalism-based density matrix reconstruction accurately predicts the mean measurement values of $x_{22}$ in the case of the statevector_simulator method for the 2-qubit SU(4) scenario. In the case of the qasm_simulator method, the proposed formalism approximately predicts the mean measurement value. This is what we expect upon measurement of the considered 2-qubit quantum states, up to statistical fluctuations. Furthermore, the IBM machine experiment’s results also show that the predicted mean measurement values of $x_{22}$ for the considered gate model are very close to the true values,

![FIG. 7. Plots of the true ($x_{22}$) versus the predicted ($x_{22}^\text{pred}$) mean measurement values for the 2-qubit SU(4) case for the shown gate model using the three backends statevector_simulator, qasm_simulator, and ibmqx2. The reconstructed density matrix accurately predicts $x_{22}$ in the case of the statevector_simulator method, whereas slight deviations are observed in the cases of the qasm_simulator method and the IBM machine experiment.](image-url)
which correspond to the mean values of $x_{22}$ obtained upon execution of the circuit on the IBM machine.

**VII. CONCLUDING REMARKS**

We discuss generating a reliable and practical inference for the quantum state of a system. We pay special reference to circumstances where information about the system not only includes the probabilities of possible outcomes but also coherences between different outcomes. The need for such an inference arises in a variety of contexts. We emphasize applications to validate the operation of noisy intermediate-scale quantum devices. There are other circumstances where the very measurement of probabilities can be measured. As we show, partial measurements can still be very useful because the inferred state can be used to compute the mean values of the missing variables.

Although our current approach is most suitable for noisy intermediate-scale quantum devices, the simplicity and predicting power of this alternative approach to quantum tomography, based on the maximal information entropy, invites generalization to other SU($N$) systems. Research is underway to generalize this approach to larger density matrices and analyze the scaling complexity of this approach to the general case of SU($N$).

**ACKNOWLEDGMENTS**

We acknowledge the financial support from the U.S. Department of Energy (Office of Basic Energy Sciences) under Award No. DE737 SC0019215. We also acknowledge use of the IBM Q for this work. The views expressed here are those of the authors and do not reflect the official policy or position of IBM or the IBM Q team.

**APPENDIX A: DETAILED RECONSTRUCTION OF THE DENSITY MATRIX FOR THE SU(4) GROUP**

Consider a quantum system that gives rise to the SU(4) Lie group algebra. To reconstruct the density matrix for such a system from a finite set of observables, we assume that $x_{11} = \langle 1 | 1 \rangle$ and $x_{12} = \langle 1 | 2 \rangle$. For the complete description of the system, we require a basis of 16 operators:

\[
\{|1 \rangle \langle 1|, |2 \rangle \langle 2|, |3 \rangle \langle 3|, |4 \rangle \langle 4|, (1|2\rangle \langle 1|2\rangle, (2|2\rangle \langle 1|2\rangle, (1|3\rangle \langle 1|3\rangle, (1|4\rangle \langle 1|4\rangle, (2|3\rangle \langle 2|3\rangle, (2|4\rangle \langle 2|4\rangle, (3|4\rangle \langle 3|4\rangle, (1|2\rangle \langle 2|1\rangle, (1|3\rangle \langle 3|1\rangle, (1|4\rangle \langle 4|1\rangle, (2|3\rangle \langle 3|2\rangle, (2|4\rangle \langle 4|2\rangle, (3|4\rangle \langle 3|4\rangle\}. \quad (A1)
\]

Based on the maximal entropy formalism we can write the density operator in terms of the operators corresponding to the available observables as
\[
\hat{\rho} = \frac{1}{Z(\lambda_{11}, \lambda_{12})} \exp(-\lambda_{11} |1\rangle \langle 1| - \lambda_{12} |2\rangle \langle 2|) \\
\times \langle 2| - \lambda_{12}^* |2\rangle \langle 1|]. \quad (A2)
\]

In this case the operators do not commute. To compute the density matrix, we first diagonalize matrix $A$, which we get upon summing the terms in the exponent within the SU(4) representation:

\[
A = \begin{bmatrix}
-\lambda_{11} & -\lambda_{12} & 0 & 0 \\
-\lambda_{12}^* & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} = \sum_{i=1}^{4} \epsilon_i |\phi_i\rangle \langle \phi_i|.
\]

The $\{\epsilon_i, |\phi_i\rangle\}$ are eigenvalues and corresponding eigenvectors of the matrix $A$, given by
\[
\epsilon_1 = 0, \quad (\phi_1) = (0, 0, 0, 1), \\
\epsilon_2 = 0, \quad (\phi_2) = (0, 0, 1, 0), \\
\epsilon_3 = -\frac{1}{2}(\lambda_{11} + \sqrt{4\lambda_{12}\lambda_{12}^* + \lambda_{11}^2}), \quad (\phi_3) = (k_3, 1, 0, 0), \\
\epsilon_4 = -\frac{1}{2}(\lambda_{11} - \sqrt{4\lambda_{12}\lambda_{12}^* + \lambda_{11}^2}), \quad (\phi_4) = (k_4, 1, 0, 0),
\]

where $k_3 = -\epsilon_3/\lambda_{12}^*$ and $k_4 = -\epsilon_4/\lambda_{12}^*$.

Now we can compute the density operator in the basis of the eigenvectors of $A$:
\[
\hat{\rho} = \frac{1}{Z} (\exp[\epsilon_1] |\phi_1\rangle \langle \phi_1| + \exp[\epsilon_2] |\phi_2\rangle \langle \phi_2| \\
+ \exp[\epsilon_3] |\phi_3\rangle \langle \phi_3| + \exp[\epsilon_4] |\phi_4\rangle \langle \phi_4|),
\]

\[
Z = \text{tr}(\exp[A]) = \sum_{i=1}^{4} \exp[\epsilon_i].
\]

As far as we computed the exponent of the operators, now we can expand the projection operators $|\phi_i\rangle \langle \phi_i|$ in our initial basis (A1):

\[
|\phi_1\rangle \langle \phi_1| = |4\rangle \langle 4|, \\
|\phi_2\rangle \langle \phi_2| = |3\rangle \langle 3|, \\
|\phi_3\rangle \langle \phi_3| = \frac{1}{\sqrt{(k_3^2 + 1)(k_3^2 + 1)}} (k_3^2 |1\rangle \langle 1| + k_3 |1\rangle \langle 2| \\
+ k_3^* |2\rangle \langle 1| + |2\rangle \langle 2|), \\
|\phi_4\rangle \langle \phi_4| = \frac{1}{\sqrt{(k_4^2 + 1)(k_4^2 + 1)}} (k_4^2 |1\rangle \langle 1| + k_4 |1\rangle \langle 2| \\
+ k_4^* |2\rangle \langle 1| + |2\rangle \langle 2|).}
\]
Giving the final form of the density operator as
\[ \hat{\rho} = \frac{1}{Z} \sum_i \exp[\varepsilon_i]|\phi_i\rangle\langle\phi_i| \]
\[ = \frac{1}{Z} \left[ |4\rangle\langle 4| + |3\rangle\langle 3| + (a + b)|1\rangle\langle 1| + \left( \frac{a}{k_3^2} + \frac{b}{k_4^2} \right) |1\rangle\langle 2| + \left( \frac{a}{k_3^2} + \frac{b}{k_4^2} \right) |2\rangle\langle 1| + \left( \frac{a}{k_3^2} + \frac{b}{k_4^2} \right) |2\rangle\langle 2| \right] \]
(A3)
where \( Z = \sum_i \exp[\varepsilon_i], k_3 = -\varepsilon_3/\lambda_{11}^*, k_4 = -\varepsilon_4/\lambda_{12}^* \),
\[ a = \frac{|k_3|^2}{\sqrt{(k_3^2 + 1)(k_3^2 + 1)}} \exp[\varepsilon_3] \]
\[ b = \frac{|k_4|^2}{\sqrt{(k_4^2 + 1)(k_4^2 + 1)}} \exp[\varepsilon_4] \]

1. Obtaining the Lagrange multipliers

Once we have the density operator defined in terms of the basis operators and the unknown Lagrange multipliers, we use the information of the measured mean values of the operators to determine the Lagrange multipliers. The method to calculate the Lagrange multipliers is straightforward. Here, we show explicitly how to calculate the Lagrange multipliers and thereby predict the unknown mean measurement value \( x_{22} \) for a 2-qubit quantum system. We consider a 2-qubit quantum circuit that is acted upon by an \( R_y \) gate (with a randomly selected rotation angle) on each qubit to obtain the mean values \( x_{11}, x_{12}, x_{22} \). The \textit{statevector_simulator} backend in Qiskit is used to simulate the quantum circuit and thereby the following mean values are obtained: \( x_{11} = 0.225\,149, x_{12} = 0.834\,282, x_{22} = 0.772\,851 \). The form of the density operator for the same is given by Eq. (A3).

Assuming that we know only \( x_{11} \) and \( x_{12} \), we use this information to solve the system of transcendental equations for the determination of unknown Lagrange multipliers \( (\lambda_{11}, \lambda_{12}) \):
\[ x_{11} = f(\lambda_{11}, \lambda_{12}) = \frac{1}{Z}(a + b), \]
\[ x_{12} = g(\lambda_{11}, \lambda_{12}) = \frac{1}{Z} \left( \frac{a}{k_3^2} + \frac{b}{k_4^2} \right) \]
(A4)
(A5)
Here \( f \) and \( g \) are functions of the unknown Lagrange multipliers and equate to the coefficients of the corresponding basis operators in the equation of the density operator (A3). Now, an efficient root finding numerical method, such as Newton’s method, Newton Raphson’s method, etc., can be employed to solve for \( \lambda_{11} \) and \( \lambda_{12} \) in Eqs. (A4) and (A5). We used the open-source PYTHON library called \textit{scipy.optimize.fsolve} to return the roots of the transcendental equations. The following values of the Lagrange multipliers are obtained for the considered example: \( \lambda_{11} = 16.800\,321, \lambda_{12} = -12.795\,453 \). To determine the unknown mean measurement value \( x_{22} \), we use the coefficient of the basis operator \( |2\rangle\langle 2| \) in Eq. (A3) and substitute the computed values of \( \lambda_{11} \) and \( \lambda_{12} \), giving
\[ x_{22} = h(\lambda_{11}, \lambda_{12}) = \frac{1}{Z} \left( \frac{a}{|k_3|^2} + \frac{b}{|k_4|^2} \right) \]
(A6)
where \( h(\lambda_{11}, \lambda_{12}) \) represents the functional form of the mean measurement value \( x_{22} \). Substituting the values of \( \lambda_{11} \) and \( \lambda_{12} \) into the above equation, we get \( x_{22} = 0.772\,851 \), which is exactly equal to the simulated value.

a. Bordered Hessian matrix

An alternative method for deriving the Lagrange multipliers is using the standard recipe based on the bordered Hessian matrix, as shown in Ref. [20]. For solving constrained optimization problems, appropriate second-order conditions in terms of the bordered Hessian matrix can be used. We consider an objective function \( f(x, y) \) whose extrema we want to find under the given constraint \( g(x, y) = c \). To solve this problem, we construct a Lagrangian function \( L = f(x, y) + \lambda [g(x, y) - c] \). The gradient of the Lagrangian function with respect to the function variables gives the unique critical point \( (x_0, y_0; \lambda_0) \). The second-order conditions can be incorporated into the bordered Hessian matrix as
\[ H = \begin{bmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{bmatrix} \]
where \( g_x \) and \( g_y \) are gradients of the function \( g(x, y) \) with respect to \( x \) and \( y \), respectively, and the \( L_{ij} \) are second-order partial derivatives of the Lagrange function. The sufficient conditions for a local extremum are
\[ |H(x_0, y_0; \lambda_0)| > 0 \quad \rightarrow \quad (x_0, y_0) \text{ is a local maximum}, \]
\[ |H(x_0, y_0; \lambda_0)| < 0 \quad \rightarrow \quad (x_0, y_0) \text{ is a local minimum}, \]
\[ H(x_0, y_0; \lambda_0) = 0 \quad \rightarrow \quad \text{no conclusion possible}. \]
In the formalism of maximum entropy, given that there are \( K \) constraints on the mean values of variables \( f_k(n) \), the general expression for the expectation values is expressed as
\[ \sum_{n=1}^{N} f_k(n) p(n) = \langle f_k(n) \rangle, \]
(A7)
where n follows the probability distribution \( p(n) \) that is constrained by the normalization of the probability distributions

\[
\sum_{n=1}^{n=N} p(n) = 1.
\]

Now, the formalism of maximum entropy results in a particular form of the probability distribution function, as proven by Jaynes [14], that satisfies the constraint equations

\[
p(n) = \frac{1}{Z} \exp \left\{ - \sum_{k=1}^{k=K} \lambda_k f_k(n) \right\}.
\]

In our proposed approach of reconstructing the density matrix, the variables whose expectation values we consider are operators \(|i\rangle\langle j|\). Therefore, in the density operator expression that corresponds to maximum entropy under the given constraints, we have operators instead of functions of variables. Thus, following the reconstruction of the density matrix, we obtain the expectation value of the operators that corresponds to \( \langle f_k(n) \rangle \) in Eq. (A7). Because of this, the construction of the bordered Hessian matrix is challenging as the functional form of the variable is the operator in our approach. However, we do satisfy the necessary conditions involved in the maximization of the entropy by choosing an appropriate form of the density operator as defined by Jaynes [14] and thereby calculating the Lagrange multipliers using the standard technique of Newton’s method or by using an inbuilt numerical solver of a PYTHON library.

2. Analysis of Lagrange multipliers

An analysis of the Lagrange multipliers is also carried out in Fig. 8 to determine the range of Lagrange multipliers that results in an accurate prediction of the unknown mean measurement \( x_{22} \). It can be seen from Fig. 8 that, when \( \lambda_{12} \) is in the range [-3, 3], then, although the maximal entropy formalism-based reconstructed density matrix accurately predicts the variation of the unknown mean measurement, the prediction is slightly deviated from the true value. However, outside this range the reconstructed density matrix very accurately predicts the value of \( x_{22} \). The slight deviation in [-3, 3] can be handled by making modifications to the unitary operator that is applied on the \( n \)-qubit quantum system.

APPENDIX B: A POSSIBLE APPROACH FOR COHERENCE MEASUREMENTS ON AN IBM QUANTUM COMPUTER

Considering the 2-qubit [SU(4)] case, it is easy to get the value of \( x_{11} = \langle 1|1 \rangle \) using measurement results from an IBM quantum computer because \( x_{11} \) corresponds to the probability of getting final results as \( |1 \rangle \). However, it is not easy to get the results of the coherence on a quantum computer, for example, \( x_{12} = \langle 1|2 \rangle \), because the operator of coherence is not available at measurements. However, this measurement can be achieved by decomposing the operator into a sum of tensor products of Pauli matrices and evaluating each tensor product on the quantum computer. Consider the simple example of evaluating \( x_{12} = \langle 1|2 \rangle \). Qiskit’s tensor product order is followed to illustrate the considered example.

1. Decomposition

We can rewrite \( |1\rangle\langle 2| \) as the matrix

\[
|1\rangle\langle 2| = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

Considering four \( 2 \times 2 \) matrices.

\[
A = \frac{1}{2}(Z + I) = \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix},
\]

\[
B = \frac{1}{2}(Z - I) = \begin{bmatrix}
0 & 0 \\
0 & 1
\end{bmatrix},
\]

\[
C = \frac{1}{2}(X + iY) = \begin{bmatrix}
0 & 1 \\
0 & 0
\end{bmatrix},
\]

\[
D = \frac{1}{2}(X - iY) = \begin{bmatrix}
0 & 0 \\
0 & 1
\end{bmatrix},
\]
where $X$, $Y$, and $Z$ are Pauli matrices and $I$ is the $2 \times 2$ identity matrix, we can get

$$
\begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} = A \otimes C
$$

or

$$
\langle 1 | 2 \rangle = A \otimes C = \frac{1}{2}(Z_2X_1 + iZ_2Y_1 + I_2X_1 + iI_2Y_1).
$$

(B1)

Thus, we have successfully decomposed the coherence operator into the tensor products of the Pauli matrices. Generally speaking, we can use the technique to decompose any coherence operator into tensor products of Pauli matrices. The evaluation of each tensor product of Pauli matrices can be achieved by using the rotation gates. More details about this technique can be found in Refs. [45,46].

---

[33] J. Yuen-Zhou, J. J. Krich, M. Mohseni, and A. Aspuru-Guzik, Quantum state and process tomography of energy


