Learning unknown pure quantum states

Sang Min Lee,[1,*] Jinhyoung Lee,[2,†] and Jeongho Bang[3,‡]
[1]*Korea Research Institute of Standards and Science, Daejeon 34113, Korea*
[2]*Department of Physics, Hanyang University, Seoul 04763, Korea*
[3]*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea*

(Received 16 May 2018; published 2 November 2018)

We propose a learning method for estimating unknown pure quantum states. The basic idea of our method is to *learn* a unitary operation $\hat{U}$ that transforms a given unknown state $|\psi_\tau\rangle$ to a known fiducial state $|f\rangle$. Then, after completion of the learning process, we can estimate and reproduce $|\psi_\tau\rangle$ based on the learned $\hat{U}$ and $|f\rangle$. To realize this idea, we cast a random-based learning algorithm, called "single-shot measurement learning," in which the learning rule is based on an intuitive and reasonable criterion: the greater the number of success (or failure), the less (or more) changes are imposed. Remarkably, the learning process occurs by means of a single-shot measurement outcome. We demonstrate that our method works effectively, i.e., the learning is completed with a *finite* number, say N, of unknown-state copies. Most surprisingly, our method allows the maximum statistical accuracy to be achieved for large N, namely $\simeq O(N^{-1})$ scales of average infidelity. It highlights a nontrivial message, that is, a random-based strategy can potentially be as accurate as other standard statistical approaches.

DOI: 10.1103/PhysRevA.98.052302

I. INTRODUCTION

The characterization of a quantum state repeatedly generated from a preparation setup is a key step for many quantum applications [1,2]. So far, these tasks have been performed with so-called "quantum state tomography (QST)" [3–5]. The conventional QST, which follows the standard statistical methodology, allows us to estimate unknown quantum states over a finite number N of registered data from a set of measurement setups optimally chosen in advance. Such a standard approach is very appealing and has been the cornerstone of these practical tasks for decades, since it appears to be likely beneficial to the extraction of information from optimized measurements. However, it was proven in Ref. [6] that one can expect at best $O(N^{-3/4})$ of infidelity in standard (local) QST, whereas $O(N^{-1})$ is the maximum accuracy on statistical grounds [7], which is imposed by the theory of quantum-state estimation [8], or equivalently, no cloning [9]. Thus achieving higher accuracy, e.g., close to $O(N^{-1})$ infidelity, is still challenging both theoretically and practically.

Recently, it has been determined that achieving a level of accuracy at least as high as in the standard QST is possible using a different strategy, namely that of changing the measurements in an *adaptive* way. In this case, the measurement setting is appropriately chosen from trial to trial depending on the previously obtained measurement outcomes [10–15]. Such adaptive QSTs have a number of practical advantages which include (i) the statistical errors are not as dominant, (ii) there is no need to deal with exponentially large data, and (iii) (post) data analysis is not required, for example, least-squares inversion [16] or maximum-likelihood correction [17]. Usually, the achievement of these advantages is established by the "optimal instructions" for the adaptive process. For example, one of the useful ways might be to use Bayesian estimation to decide the next-stage measurements [12,18]. Quite recently, a variant of such adaptive strategies, called self-guided QST, has been proposed with improved accuracy and efficiency [19,20].

In this paper, we propose a method of machine learning to estimate unknown pure quantum states. The main idea of our method is to *learn* a unitary operation $\hat{U}$ that transforms a given unknown state $|\psi_\tau\rangle$ to a known fiducial state $|f\rangle$. Then, after the learning is completed, we can infer and reproduce the unknown state $|\psi_\tau\rangle$ such that $|\psi_\tau\rangle \simeq |\psi_{\tau,\mathrm{est}}\rangle = \hat{U}^\dagger|f\rangle$. To do this, we employ a learning algorithm, called "single-shot measurement learning (SSML)" [21,22]. Our SSML method runs based on the single-shot measurement outcomes (as casted in its name), which is the most distinctive feature compared to other adaptive proposals. Owing to this feature, we additionally expect little requirement of classical computational resources: for example, additional calculations for assessing the performances are not required at each learning step. Furthermore, as our method is also akin to an adaptive approach, the practical advantages (i)–(iii) can be achieved by invoking the adaptivity. Most importantly, our scheme shows that the average infidelity $\bar{\varepsilon} = 1 - \int d\psi_\tau |\langle\psi_{\tau,\mathrm{est}}|\psi_\tau\rangle|^2$ scales $\simeq O(N^{-1})$. This result is comparable to the yields from the standard QST in the case where our implicit assumption—i.e., the unknown state is pure—is brought. This indicates that our SSML is able to be as efficient as the standard QST for pure-state estimation.

II. SCHEME AND METHOD

We briefly describe how our method proceeds by specifying the key elements. First, let us consider a preparation device

*samini@kriss.re.kr
†hyoung@hanyang.ac.kr
‡jbang@kias.re.kr

(**P**) which can repeatedly generate unknown pure state $|\psi_\tau\rangle$. Here, it is assumed that **P** is initially given and the Hilbert-space dimension d of $|\psi_\tau\rangle$ is also known. We assume further that it is impractical to take **P** to pieces for the purpose of the state estimation. Excepting some very specific problems, such assumptions have commonly been made. We also set a part of operation device (**U**) for the implementation of an arbitrary unitary $\hat{U}(\mathbf{p})$, where $\mathbf{p}$ is the vector whose components are *controllable* learning parameters. We then choose a fiducial state $|f\rangle$ freely, and let the measurement device (**M**) correspond to a "yes-or-no" question, namely of whether we get the desired target:

$$\hat{M}_f = |f\rangle\langle f|, \quad \hat{M}_{f^\perp} = \hat{\mathbb{1}} - \hat{M}_f. \tag{1}$$

For convenience, the Hilbert-space dimension of $|f\rangle$ is assumed to be equal to d. Then, by connecting these three elements, we can define a system of the learning building block, i.e., **P**-**U**-**M** (the "student," say), for conventional quantum information processing. In such a setting, we employ another key element which is the feedback system (**F**). It is responsible for the training (the "teacher," say). **F** has an optimal learning algorithm and a relatively small size of the (classical) memory to record the learning parameters. Then, the goal of the learning is to find a learning parameter vector $\mathbf{p}_{\rm est}$ close to an optimal one in $\{\mathbf{p}_{\rm opt}\}$, and finally estimate as

$$\underbrace{|\psi_\tau\rangle = \hat{U}(\mathbf{p}_{\rm opt})^\dagger |f\rangle}_{\text{actual unknown state}} \simeq \underbrace{|\psi_{\tau,\rm est}\rangle = \hat{U}(\mathbf{p}_{\rm est})^\dagger |f\rangle}_{\text{estimated state}}. \tag{2}$$

Here, we note that the presented method can be referred to as a *quantum-classical hybrid* learning concept; i.e., the student is quantum and the teacher is classical. Such a hybridization would be easier and more economical to realize. There is also the possibility of gaining a quantum advantage from the quantum student [23,24].

III. SINGLE-SHOT MEASUREMENT LEARNING

The efficiency and accuracy of our method strongly depends on the learning algorithm. Here we employ a learning algorithm, called "single-shot measurement learning (SSML)" [25]. The most intriguing feature of the SSML is that the learner (i.e., $\hat{U}$ here) updates its own parameters by means of the single-shot measurement outcomes [26]. Specifically, the SSML runs as follows: for every learning step n, **P** generates $|\psi_\tau\rangle$ and it is transformed to an output state through **U**. Then, **M** performs the projective measurement with $\{\hat{M}_f, \hat{M}_{f^\perp}\}$, where each outcome is identified as a "success" or a "failure." More specifically, if a measurement result is $|f\rangle$, this is a success and regarded as one successful trial of the target task. Otherwise, we have a fail outcome. Thus we can infer that if the learning proceeds as expected, **M** will produce the more success outcomes; i.e., the number of *consecutive* successes, denoted $M_S^{(n)}$, can be regarded as an index of how close the control parameters in $\mathbf{p}^{(n)}$ at the current n step are to an optimal value $\in \{\mathbf{p}_{\rm opt}\}$. As such, the rule for updating $\mathbf{p}$ is made as below.

[**R.1**] When we get a success outcome, **F** follows

$$M_S^{(n)} \leftarrow M_S^{(n-1)} + 1, \quad \mathbf{p}^{(n+1)} \leftarrow \mathbf{p}^{(n)}. \tag{3}$$

At the first step, i.e., for $n = 1$, we set $\mathbf{p}^{(1)} \leftarrow \mathbf{r}$ and $M_S^{(0)} \leftarrow 0$, where $\mathbf{r}$ is a random vector whose components consist of random numbers.

[**R.2**] Otherwise, if the outcome is fail, **F** proceeds as

$$M_S^{(n)} \leftarrow 0, \quad \mathbf{p}^{(n+1)} \leftarrow \mathbf{p}^{(n)} + \omega\mathbf{r}, \tag{4}$$

where $\omega = \alpha(M_S^{(n-1)} + 1)^{-\beta}$ is the weight for the random vector $\mathbf{r}$. Here, α and β are the free parameters related to the algorithm's performance.

Note that adopting the random vector $\mathbf{r}$ in [**R.2**], instead of using a preprogrammed one, is a typical strategy of machine learning [27], and is of particular importance, since it is expected to alleviate the local-minimum problem [28]. These learning rules of the SSML—i.e., the greater the number of success (fail), the less (more) changes are imposed—intuitively makes sense.

The learning is not completed until $M_S^{(n)}$ becomes sufficiently large while producing no fail; more specifically, the learning is completed when the condition $M_S^{(n)} = M_H$ is met. We call this the "halting condition." After the learning is completed by satisfying this halting condition, we can obtain $\hat{U}(\mathbf{p}_{\rm est})$ with $\mathbf{p}_{\rm est} \leftarrow \mathbf{p}^{(n)}$. Here, the total iteration n is the consumption N of state copies for the estimation in Eq. (2). The learned $\hat{U}(\mathbf{p}_{\rm est})$ is then expected to transform $|\psi_\tau\rangle$ to $|f\rangle$ faithfully, i.e., satisfying the following condition (for $M_H \gg 1$):

$$\varepsilon = 1 - |\langle f|\hat{U}(\mathbf{p}_{\rm est})|\psi_\tau\rangle|^2 = 1 - |\langle\psi_{\tau,\rm est}|\psi_\tau\rangle|^2 \ll 1. \tag{5}$$

Note here that there exists a trade-off relation between inaccuracy and the learning time, depending on the predetermined number M_H; the larger (smaller) M_H, the lower (higher) infidelity ε we have and the more (less) iterations or equivalently unknown-state copies in our case are required to complete the learning process. Thus it is very important to choose appropriate M_H to account for the desired learning accuracy and time.

IV. QUBIT-STATE ESTIMATION

To analyze our method, we here consider the estimation of unknown single-qubit state. Considering the possible realization of our approach, we adopt a general unitary learner, parametrized as

$$\hat{U}(\mathbf{p}) = \exp(-i\,\mathbf{p}^T\mathbf{G}), \tag{6}$$

where $\mathbf{p} = (p_x, p_y, p_z)^T$ is the control parameter vector and $\mathbf{G} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)^T$ is an operator vector whose components are SU(2) generators, i.e., Pauli operators. Note that p_j ($j = x, y, z$) can be represented by real hands-on control parameters, e.g., wave-plate angles for a polarization qubit in a linear-optical setup (see Appendix A 1).

A. Estimation of pure state

First, we investigate whether our SSML method works well, i.e., whether the learning is completed in finite learning steps. To do this, we need to introduce the learning probabilities $P(N)$ defined as the probability that the learning is completed before or at a number N of learning iterations [23]. Remarkably, the learning probability $P(N)$ is here analyzed

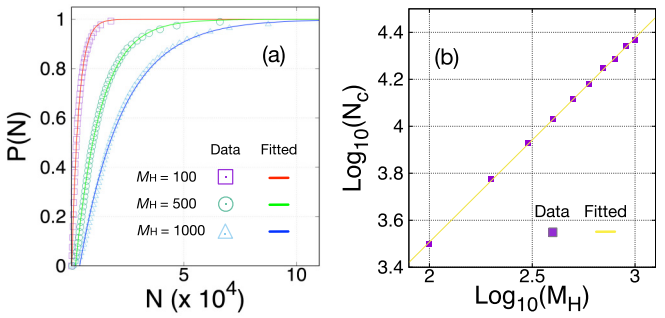

FIG. 1. (a) Learning probabilities $P(N)$ are drawn for $M_H = 100$, 500, and 1000. For each M_H, the data are obtained from 10^4 trials. In each trial, $|\psi_\tau\rangle$ is made at random. The data are well fitted to $1-\exp(-N/N_c)$ with the factor N_c, which means the required number of state copies for the completion of the learning. We get $N_c \simeq 3158$ for $M_H = 100$, $\simeq 13037$ for $M_H = 500$, and $\simeq 23377$ for $M_H = 1000$. (b) The graph of N_c versus M_H is also given on the $\log_{10}$-$\log_{10}$ scale. The simulations are performed by increasing M_H from 100 to 1000 at intervals of 100. Each data point of N_c is obtained from 10^4 simulations. By fitting the data, we get $N_c \simeq O(M_H^{0.869})$ (the detailed data are listed in Appendix A 4).

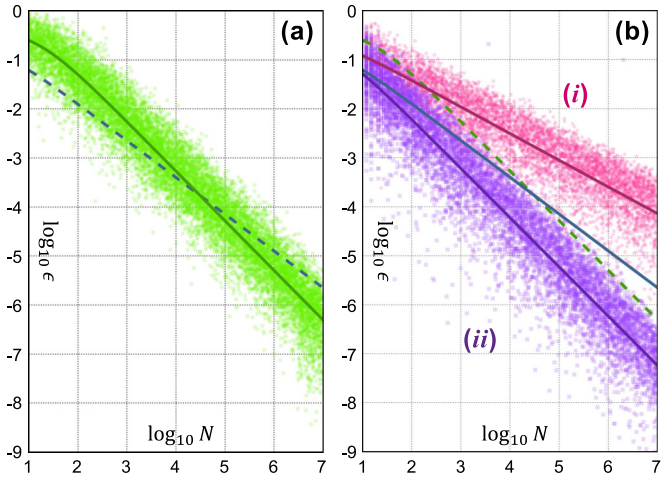

FIG. 2. Infidelities ε are evaluated from (a) SSML and (b) SQST. We depict the graph of ε versus N as dots and their fitting lines on a $\log_{10}$-$\log_{10}$ scale. The SSML result, i.e., $\overline{\varepsilon}_{\mathrm{SSML}}$, exhibits the ultimate statistical accuracy, $O(N^{-\gamma})$ with $\gamma \simeq 1$ (solid green line). In the case of the SQST, the data are sorted into the two groups, each of which is characterized from those of the mixed and pure estimating states. The former group exhibits $O(N^{-\frac{1}{2}})$ (upper red line) and the latter does $O(N^{-1})$ (lower purple line) (for detail, see the main manuscript). We note that we draw the (dashed and same colored) fitting lines in opponents for clearer comparison.

as $\simeq 1-\exp(-N/N_c)$ with a *finite constant* N_c. Note that, because $P(N)$ is a cumulative distribution by definition, the constant N_c denotes the average iterations required to complete the learning (for detailed analysis, see Appendix B). Thus the result means that, in most cases, learning is expected to be completed within certain (i.e., N_c) learning steps. To verify this prediction, we performed numerical simulations: 10^4 trials for each different halting condition M_H. Here, α and β in Eq. (4) are chosen to maximize the learning efficiency: $\alpha = 0.3$ and $\beta = 0.5$ (see Appendix A 3). We hereby note that the simulation is carried out, considering the linear-optical realization (see Appendixes A 1 and A 2). The unknown states $|\psi_\tau\rangle$ are also randomly chosen for each trial. We extract the learning probabilities $P(N)$ from the obtained data and show that they are well fitted to the aforementioned function $1-\exp(-N/N_c)$. Here, N_c is estimated as $\simeq O(M_H^{0.869})$ (see Fig. 1). The results are in good agreement with our theoretical predictions.

Secondly, we investigate the accuracy: i.e., the average infidelity $\overline{\varepsilon} = 1 - \int d\psi_\tau |\langle\psi_{\tau,\mathrm{est}}|\psi_\tau\rangle|^2$ for large N. Noting that the adaptive estimators can be precise in a metrological scenario [15,24,29–31], we expect that our SSML can exhibit similar behavior. To corroborate this, simulations are performed. The data from the standard QST (SQST) are also analyzed for comparison, where the observables are chosen from $\{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ on each qubit. Figure 2 represents the results of our simulation in the form of ε versus N graphs on the $\log_{10}$-$\log_{10}$ scale. By fitting the obtained data to $\varepsilon = C(N + N_0)^{-\gamma}$ [32], we evaluate the average infidelity $\overline{\varepsilon}$ with the main factor γ, such that $\overline{\varepsilon} \simeq O(N^{-\gamma})$. Here, in the SSML, we can achieve that $\overline{\varepsilon}_{\mathrm{SSML}} \simeq O(N^{-1})$. In the case of SQST, the final results are sorted into the two groups. (i) About half of the data (denoted by the red pentagon points) are reconstructed to the mixed states (i.e., inside the Bloch sphere, geometrically) without the maximum-likelihood correction. (ii) On the other hand, the other half of the data (denoted by the purple square points) are initially not the legitimate physical states (i.e., outside the Bloch sphere). Therefore, we need to correct the data using the maximum-likelihood correction, so that they are transformed to the pure state (i.e., on the Bloch sphere). Thus the infidelities evaluated from (i) are higher than those from (ii) in general. Fitting the data corresponding to each case (i) and (ii), we can find that $\overline{\varepsilon} \simeq O(N^{-\frac{1}{2}})$ (upper red line) and $\overline{\varepsilon} \simeq O(N^{-1})$ (lower purple line). However, we cannot sort the data groups in an actual situation and may observe only $\overline{\varepsilon} \simeq O(N^{-\frac{3}{4}})$ (middle blue line) for overall data in a half-half manner. Nevertheless, if we use the additional information—i.e., the fact that the unknown state is pure which is implicitly assumed in SSML—then, the maximum-likelihood correction moves all states to the pure state [i.e., the data in (i) are transferred into (ii)] and the overall data are equal to the case of (ii); namely, we get $\overline{\varepsilon} \simeq O(N^{-1})$ (lower purple line). The above results support the idea that a learning estimation based on a random strategy is able to be efficient and accurate (see also Refs. [7,11,19,20]).

B. Estimation of mixed state close to pure

Analyzing further, we consider the learning of mixed states, even though our method is developed for estimating pure states. Normally, pure states are of more interest, because they carry the full quantum information and many practical results are derived for the pure states. However, as a faulty apparatus produces the impurity on the states, we need to have the strategies to learn the mixed state still close to pure [33]. In fact, our original SSML method is optimal for finding the direction of the Bloch vector of the unknown state rather than

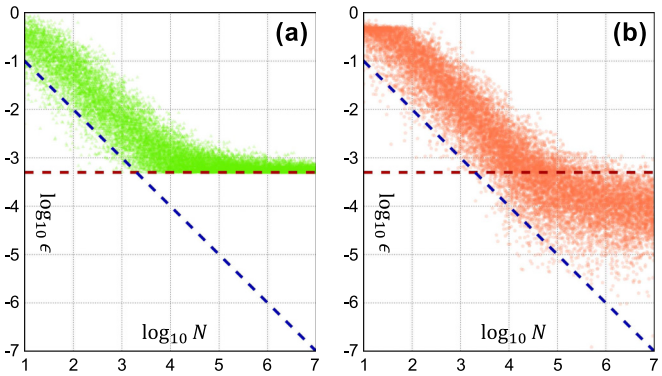

FIG. 3. Results of the simulations for the unknown mixed state $\hat{\rho}$ in Eq. (7). We plot graph of ε versus N, by adopting (a) a primitive strategy, i.e., without any change of the protocol and (b) slightly developed strategy. The simulations are performed for $q = 10^{-3}$. The infidelity ε is shown to be bounded by Eq. (8) in the first strategy, while it decreases continuously with consuming the state copies below the bound in our second strategy. In both strategies, the tendency of $\overline{\varepsilon} \simeq O(N^{-1})$ can still be exhibited above the bound.

its size, namely, the purity [34]. Nevertheless, our SSML is applicable to the learning *nearly pure* states. To show this, we consider an illustrative example state, assuming the isotropic noises, as

$$\hat{\rho}_\tau = (1-q)|\psi_\tau\rangle\langle\psi_\tau| + q\,\frac{\hat{\mathbb{1}}}{2}, \tag{7}$$

where the fraction of the noise q is assumed to be small.

Here, we consider two strategies. The first one is a primitive way, that is to apply the SSML without any change. The infidelity ε is then predicted to be bounded by

$$1 - \langle\psi_\tau|\hat{\rho}_\tau|\psi_\tau\rangle = \frac{q}{2}. \tag{8}$$

We perform the simulations for $q = 10^{-3}$ to corroborate our prediction. The result is given in Fig. 3(a), where the infidelity can never be smaller than the bound $\frac{q}{2}$ (dotted horizontal red line) as predicted, even though the tendency of $\overline{\varepsilon} \simeq O(N^{-1})$ is still exhibited above the bound.

The problem of the infidelity bound can be alleviated by our second strategy. At first, we note that the hating condition is not available for the mixed-state learning because M_S could not reach up to M_H with the impurity of the unknown state. Thus we adopt the strategy to memorize the maximum of success events, say $M_{S,\max}$, and the parameter vector $\mathbf{p}$ at the moment; nevertheless, the learning rules [**R.1**] and [**R.2**] are still applicable. Then, after the completion of the learning, additional $M_{S,\max}$ of measurements are performed on $\mathbf{p}_{\rm est}$. The final state is estimated as

$$\frac{m_f}{M_{S,\max}}|\psi_{\tau,\rm est}\rangle\langle\psi_{\tau,\rm est}| + \frac{m_{f^\perp}}{M_{S,\max}}|\psi^\perp_{\tau,\rm est}\rangle\langle\psi^\perp_{\tau,\rm est}|, \tag{9}$$

where m_f is the number of counts of the final measurement $\hat{M}_f = |f\rangle\langle f|$ and $m_{f^\perp} = M_{S,\max} - m_f$. We also perform the simulations for $q = 10^{-3}$. In Fig. 3(b), we give the result, where the tendency $\overline{\varepsilon} \simeq O(N^{-1})$ is observed until ε reaches to the bound $\frac{q}{2}$; however, the accuracy is improved continuously below the bound (for more detailed data, see Appendix A 6).

V. SUMMARY AND REMARKS

We have presented a simple but powerful method to estimate unknown pure quantum states $|\psi_\tau\rangle$. The main idea was to *learn* a unitary $\hat{U}$ to perform $|\psi_\tau\rangle \to |f\rangle$ for a known fiducial state $|f\rangle$. Then we could estimate $|\psi_\tau\rangle$, such that $|\psi_\tau\rangle \simeq |\psi_{\tau,\rm est}\rangle = \hat{U}^\dagger|f\rangle$. To realize this idea, we casted a learning algorithm, called single-shot measurement learning (SSML), in which the learner ($\hat{U}$ here) was renewed according to a reasonable learning rule, i.e., the greater the number of success (fail), the less (more) adjustment is imposed. We noted that basically our method can be understood as a (weighted) random learning process with one-by-one measurements. As our approach is akin to the other adaptive approaches, the advantageous features from the "adaptivity" can be carried over. Most importantly, we demonstrated that our method works well for a finite number of state copies. Most surprisingly, we obtained nearly $O(N^{-1})$ level of average infidelity. This result implies an important and nontrivial scientific message, i.e., a random estimator can exhibit high accuracy in quantum state estimation.

Our method is also associated with a particular operational advantage in that the estimated state $|\psi_{\tau,\rm est}\rangle$ can *directly* be reproduced even with no identification of the found parameters in $\mathbf{p}_{\rm est}$. For example, consider a quantum linear-optical system (e.g., see Fig. 4 in Appendix A 1), in which some systematic errors, such as nonideal phase retardation of the wave plates, are inevitable [35]. Such errors may have an influence on the final estimation in theory. However, our method is not affected by such an error; specifically, if we replace the measurement part (e.g., SPD for "success" side in Fig. 4) to a single-photon source, it naturally becomes the setting for the preparation of $|\psi_{\tau,\rm est}\rangle$. Thus we do not need to reconstruct the preparation setting with the identified (may be "poor") parameters. Such an advantage cannot be found in SQST and other existing AQST schemes, since those methods implicitly assume that the theoretical description and the implementation of the experiment are perfectly matched. Actually, this advantage is of particular significance, e.g., in quantum cryptographic scenarios (see Refs. [36,37]).

We believe that our SSML method will find immediate application in quantum information tasks requiring (nearly) pure-state estimation.

ACKNOWLEDGMENTS

We are grateful to Jaewan Kim and Marcin Wieśniak for helpful discussions. J.B. was supported by the research project on quantum machine learning (No. 2018-104) of the ETRI affiliated research institute. S.M.L. and J.B. acknowledge the support of the R&D Convergence program of NST (National Research Council of Science and Technology) of Republic of Korea (No. CAP-18-08-KRISS). S.M.L. was also supported by KRISS projects (No. KRISS-2018-GP2018-0012, -0017). J.L. acknowledges the financial support of the Basic Science Research Program through the National Research Foundation of Korea (NRF) grant (No. 2014R1A2A1A10050117).

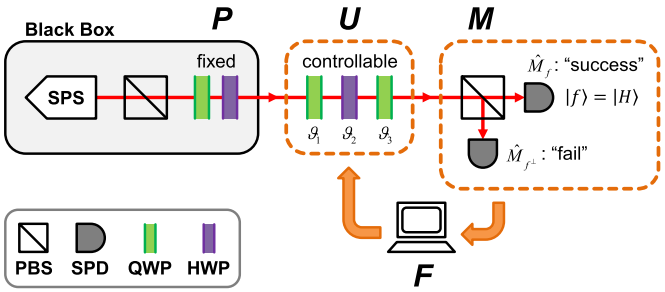

FIG. 4. Schematic layout for a linear-optical implementation of our SSML estimation.

APPENDIX A: DETAILS ON THE THEORETICAL AND NUMERICAL ANALYSIS OF SSML

1. Polarization-based linear-optical experimental setting

To demonstrate that our method works well *in a real experiment* and with a high accuracy, we performed numerical simulations, particularly considering a polarization-based linear-optical realization (see Fig. 4). First, we consider a single-photon source (SPS) and a combination of a quarter or half wave plate (QWP or HWP) to construct **P**, by assuming that these elements are completely capsulated, e.g., in a black box. Thus the single-photon state, $a|H\rangle + b|V\rangle$, generated in **P** is assumed to be an unknown state $|\psi_\tau\rangle$ (here, $|H\rangle$ and $|V\rangle$ denote state of the horizontally and vertically polarized single photon, respectively). Then, we employed a finite number of *controllable* wave plates to implement **U**; i.e., the combination of QWP(ϑ_1)-HWP(ϑ_2)-QWP(ϑ_3), where the rotation angles ϑ_i $(i = 1, 2, 3)$ consist of the control parameter vector $\mathbf{p} = (\vartheta_1, \vartheta_2, \vartheta_3)^T$. These parameters ϑ_i $(i = 1, 2, 3)$ are replaced by the general parameter vector, defined in Eq. (6) of the main manuscript. Here we note that, if the fiducial state $|f\rangle$, or equivalently, the measurement $\hat{M}_f = |f\rangle\langle f|$, is fixed, $\hat{U}$ can be considered as an internal operation of a projective measurement $\hat{U}|f\rangle\langle f|\hat{U}^\dagger$. Then, our scheme might be able to be rather simplified reducing the number of control parameters; for example, SSML works by controlling two wave plates, i.e., the ideal set of QWP and HWP for $|f\rangle = a|H\rangle + b|V\rangle$ with real a and b. However, we indicate that $|f\rangle$ can be chosen favorably, so that methodologically we need to develop our scheme for an arbitrary measurement $\hat{M}_f$; thus the total number of control parameters are to be three, which is required to characterize a unitary operation for qubit. The wave-plate combination is the minimal requirement for an arbitrary single-qubit unitary operation [38]. The measurement **M** is implemented with the polarization beam splitter (PBS) and two single-photon detectors (SPDs). Here, for the sake of simplicity, we set $|H\rangle$ to be the fiducial state.

2. Number of wave plate for U

The implementation of **U** can be performed by various combinations of wave plates. Even though the minimum requirements of the wave plate for an arbitrary unitary operation is QWP-HWP-QWP (QHQ), at least in theory the combination of QH can convert the *fixed* fiducial state $|H\rangle$ to an arbitrary state, as described in the previous subsection. We thus investigate three types of wave-plate combinations: (a) QH, (b) QHQ, and (c) QHQH. In the investigation, it is found that, theoretically, we can achieve $\gamma \simeq 1$ for all three cases with $\alpha \simeq 0.3$ and $\beta \simeq 0.5$ (see Fig. 5). Note, however, that the amount of phase retardation of the wave plate is not ideal in practice and in this case QH would not allow a general transform (for details, see Sec. 3.1.2 in Ref. [39]). However, note further that, as the number of wave plates increases, such imperfections will accumulate. Thus we utilized a QHQ setting in this work.

3. Optimization of the SSML parameters α and β

In order to optimize the feedback range ω in Eq. (4) of the main text, we examine the relationship between ε and M_S. To do this, let us assume that $\mathbf{p}$ is near to an optimal $\mathbf{p}_{\rm opt}$, but still not sufficient to complete the algorithm. Noting that $F = |\langle f|\hat{U}(\mathbf{p})|\psi_\tau\rangle|^2$, the probability that we get the number M_S of successes continuously is given as $p(M_S) = F^{M_S}(1-F)$. Thus we have

$$\overline{M_S} = \frac{F}{1-F} \simeq \frac{1}{1-F} = \frac{1}{\varepsilon}. \tag{A1}$$

Because F is close to 1 with $\mathbf{p} \simeq \mathbf{p}_{\rm opt}$, the infidelity ε is approximated as $\simeq A(\mathbf{p} - \mathbf{p}_{\rm opt})^2$. Then, the distance between

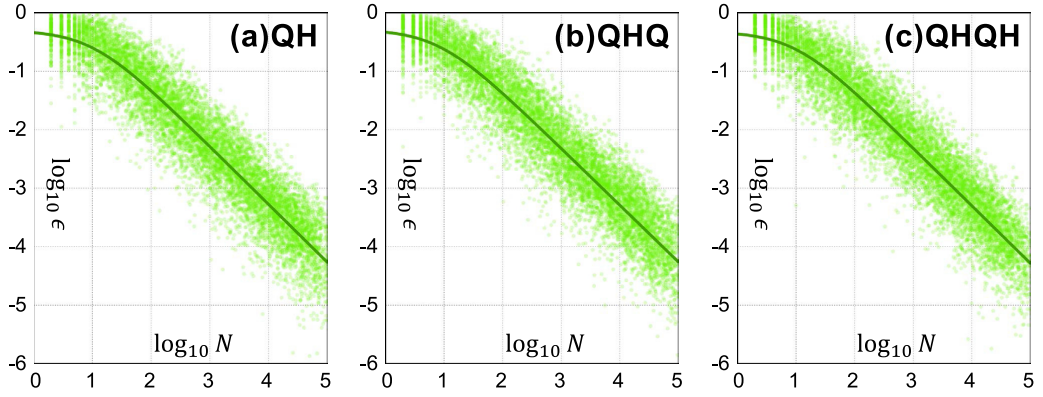

FIG. 5. Graphs of $\log_{10} N$ versus $\log_{10}\varepsilon$ for (a) QH, (b) QHQ, and (c) QHQH wave-plate combinations.

TABLE I. Detailed values of the fitting parameter N_c are listed for each M_H. The values $\overline{N}_{\text{data}}$ are obtained by averaging over 10^4 data of simulations.

M_H	N_c ($\overline{N}_{\text{data}}$)	M_H	N_c ($\overline{N}_{\text{data}}$)
100	≃3158 (≃3354)	600	≃15162 (≃15240)
200	≃5942 (≃6096)	700	≃17657 (≃17777)
300	≃8513 (≃8780)	800	≃19327 (≃19464)
400	≃10692 (≃10951)	900	≃21915 (≃22112)
500	≃13037 (≃13255)	1000	≃23377 (≃23381)

$\mathbf{p}$ and $\mathbf{p}_{\text{opt}}$ is represented as

$$|\mathbf{p} - \mathbf{p}_{\text{opt}}| \propto \overline{M_S}^{-1/2}. \tag{A2}$$

The parameter β is thus approximately 0.5, which is in good agreement with the simulation results. However, α should be found in a heuristic manner. In this case, we found that the optimal setting is $\alpha \simeq 0.3$.

4. Detailed simulation results of the learning probability

We describe detailed simulation results of the learning probability in this subsection. In particular, in order to investigate whether or not the learning is completed in a finite number of learning steps as predicted in the previous subsection, we analyze the learning probability $P(N)$, which is defined as the probability that the learning is completed before or at a number N of learning iterations. Here, let us recall the fact that because $P(N)$ is a cumulative distribution, the constant factor N_c in Eq. (B7) can be interpreted as the average number of iterations for the completion of the learning. Having the aforementioned in mind, we perform numerical simulations for analysis. The simulations are performed for different halting conditions (from $M_H = 100$ to $M_H = 1000$ at intervals of 100, total 10 cases). For each case of M_H, we perform 10^4 simulations to construct the learning probability $P(N)$. By fitting the obtained data to the function in Eq. (B7), we identified the characteristic factors N_c (see Table I). The N_c values are well matched to the average iterations $\overline{N}_{\text{data}}$ evaluated from the actual simulation data. Our analytical predictions in the previous subsection are thus well borne out.

5. More simulation data in the mixed-state learning

Here we present more detailed results of the simulations for the unknown mixed state $\hat{\rho}_\tau$ in Eq. (7),

$$\hat{\rho}_\tau = (1-q)|\psi_\tau\rangle\langle\psi_\tau| + q\,\frac{\hat{\mathbb{1}}}{2}, \tag{A3}$$

where the fraction of the noise q was assumed to be small. Based on the two strategies described in our main text, the simulations were carried out for (a) $q = 10^{-1}$, (b) $q = 10^{-2}$, and (c) $q = 10^{-3}$. Here, we also consider the case of (d) $q = 0$ for comparison. The results are given in Fig. 6 and Fig. 7. Here, it is seen that ε is bounded by Eq. (8) in the former strategy, while in the later case the accuracy can continuously be improved even when ε reaches to the bound. The behavior of $\overline{\varepsilon} \simeq O(N^{-1})$ is observed above the bound.

APPENDIX B: LEARNING PROBABILITY AND THE EFFECTIVENESS OF SSML

Here we approximately estimate $P(N)$ by using the random learning strategy, which is often casted for the analysis. To this end, we first consider the probability $p_s = (1-\varepsilon)^{M_H}$ that the learning is completed for $\mathbf{p}_{\text{est}} \simeq \mathbf{p}_{\text{opt}}$. Here, $1-\varepsilon$ is the probability of the success event, namely that of measuring the fiducial state $|f\rangle$. Then, we introduce a continuous function,

$$\tfrac{1}{2} \leqslant \Xi(\mathbf{p}) = \xi_1(p_1)\xi_2(p_2)\xi_2(p_3) \leqslant 1, \tag{B1}$$

satisfying $\Xi(\mathbf{p} \neq \mathbf{p}_{\text{opt}}) < \Xi(\mathbf{p}_{\text{opt}}) = 1$. We note that this function $\Xi(\mathbf{p})$ is obtained by minimizing

$$\left| p_s^{1/M_H} - \Xi(\mathbf{p}_{\text{opt}}) \right|. \tag{B2}$$

Thus we can assume that, *for very large* M_H,

$$p_s = (1-\varepsilon)^{M_H} \simeq \Xi(\mathbf{p}_{\text{est}})^{M_H} \simeq \Xi(\mathbf{p}_{\text{opt}})^{M_H}. \tag{B3}$$

We then adopt an interesting idea, by approximating $\xi_j(p_{j,\text{est}})^{M_H}$ $(j = 1, 2, 3)$ with a delta function,

$$\xi_j(p_{j,\text{est}})^{M_H} \approx \exp\left[-\frac{(p_{j,\text{est}} - p_{j,\text{opt}})^2}{\sigma^2} \right]. \tag{B4}$$

Actually, such an approximation is true when the control parameter space is homogeneous and isotropic. Thus Eq. (B4) is valid only for the area very close to the solution. However,

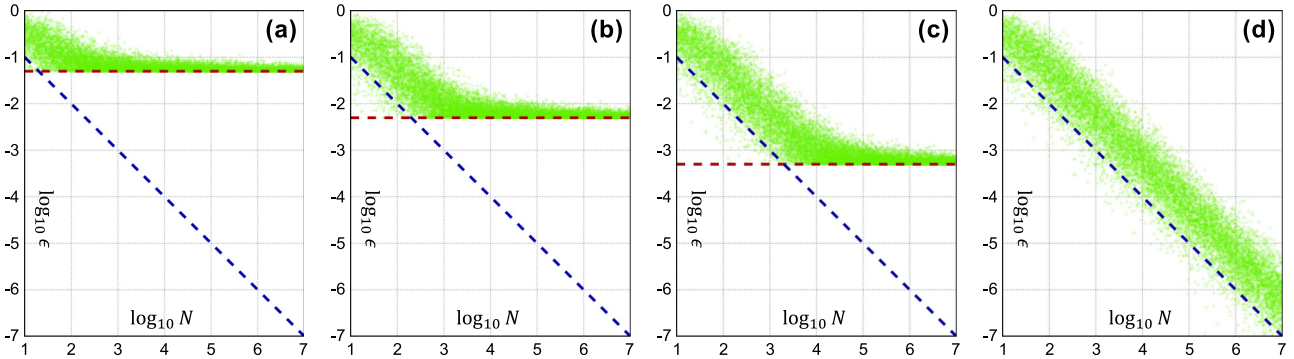

FIG. 6. Result of the simulations for the unknown mixed state $\hat{\rho}$ in Eq. (7). We plot graph of ε versus N, by adopting a primitive strategy, i.e., without any change of the protocol. The simulations are performed for (a) $q = 10^{-1}$, (b) $q = 10^{-2}$, (c) $q = 10^{-3}$, and (d) $q = 0$. The results show that the tendency of $\overline{\varepsilon} \simeq O(N^{-1})$ can be exhibited, but bounded by the degree of q as in Eq. (8).

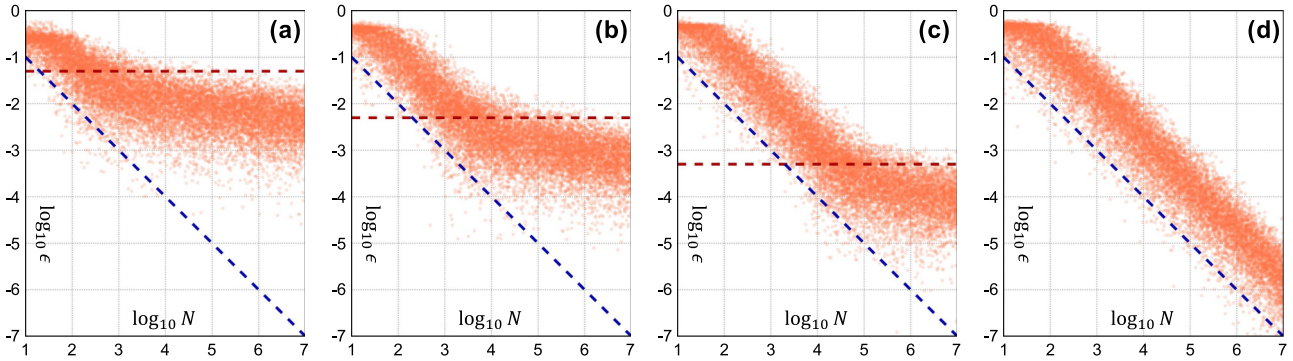

FIG. 7. Simulation results obtained by our modified strategy are given. Here, we also consider the estimation of $\hat{\rho}$ in Eq. (7). The graph of ε versus N is depicted as dots for the cases of (a) $q = 10^{-1}$, (b) $q = 10^{-2}$, (c) $q = 10^{-3}$, and (d) $q = 0$. The infidelity ε is shown to decrease with consuming the state copies, even after the bound in Eq. (8).

by observing the learning behavior, we found that the most resources for learning is used up for finding the best optimized solution near the solution points, rather than for increasing the fidelity in the starting stages. Thus our approximated error model in Eq. (B4) is expected to be reasonable. As such, we estimate the average probability $\overline{p_s}$, such that (for $\sigma \ll 1$)

$$\overline{p_s} \simeq \prod_{j=1,2,3} \int dp\, \xi_j (p_{j,\mathrm{est}})^{M_H} \approx (\sqrt{\pi}\sigma)^3, \tag{B5}$$

where σ is the value of the deviation of $p_{j,\mathrm{est}}$ about the optimal $p_{j,\mathrm{opt}}$. Actually, Δ would be vanishingly small for $M_H \gg 1$. The integration limits, from $-\infty$ to ∞, are approximated by this condition. Here, the approximation of the right part of Eq (B5) is made with the assumption that the space of $(p_{1,\mathrm{est}}, p_{2,\mathrm{est}}, p_{3,\mathrm{est}})$ is isotropic [40]. Then, for any sequence $\mathbf{p}^{(0)} \to \mathbf{p}^{(1)} \to \mathbf{p}^{(2)} \to \cdots \to \mathbf{p}^{(N)} = \mathbf{p}_{\mathrm{est}} \simeq \mathbf{p}_{\mathrm{opt}}$ of updating the parameter vectors in the learning process, we can approximate the learning probability $P(N)$ as

$$P(N) \approx \sum_{k=0}^{N-1} (1 - \overline{\Xi(\mathbf{p}^{(k)})^{M_H}})^k \overline{\Xi(\mathbf{p}^{(N)})^{M_H}}. \tag{B6}$$

Here, let us assume that the learning process is started with a parameter vector $\mathbf{p}^{(0)}$ close to $\mathbf{p}_{\mathrm{opt}}$. This assumption is reasonable, since almost all iterations would go on finding the solution near to, rather than far from, $\mathbf{p}_{\mathrm{opt}}$ (as will be seen in our numerical simulation). Then we have $\overline{p_s} \simeq \overline{\Xi(\mathbf{p}^{(j)})^{M_H}}$ for all $j = 0, 1, \ldots, N$, and finally arrive at

$$P(N) \approx 1 - (1 - \overline{p_s})^N = 1 - e^{-\frac{N}{N_c}}, \tag{B7}$$

for very large N. Here, $N_c \simeq \overline{p_s}^{-1} \simeq (\sqrt{\pi}\sigma)^{-3}$, which is the average number of iterations to complete the learning. This also indicates that we need a large iteration to achieve more accurate learning.

APPENDIX C: ENSEMBLE-BASED LEARNING

Using Eq. (A1), we can determine the updating range $\omega = \alpha(M_S + 1)^{-\beta}$. However, since $\Delta M_S = \sqrt{F}/(1 - F)$ obtained from $p(M_S)$ is very large when $F \simeq 1$, the determined value of ω can be considered to be unreliable. Thus one can consider the learning via an ensemble measurement that deals with a number M_E of samples, instead of a single shot under the same setup $\mathbf{p}^{(n)}$. In particular, such an ensemble-based learning can be considered to be more accurate and efficient than the single shot, as the fluctuation ΔM_S is very small when $F \simeq 1$. For the ensemble measurements, M_S represents the total number of successes in the single parameter $\mathbf{p}^{(n)}$, rather than consecutive successes. To test the ensemble-based measurement learning (EML), we perform numerical simulations according to the following rules. First, M_E copies are measured at nth learning step. Then, $\mathbf{F}$ updates $\mathbf{p}^{(n)}$ by using the number $M_S^{(n)}$ of measurement results, such that

$$\mathbf{p}^{(n+1)} \leftarrow \mathbf{p}^{(n)} + \alpha \left(\frac{M_E - M_S^{(n)}}{M_E} \right)^{\beta} \mathbf{r}, \tag{C1}$$

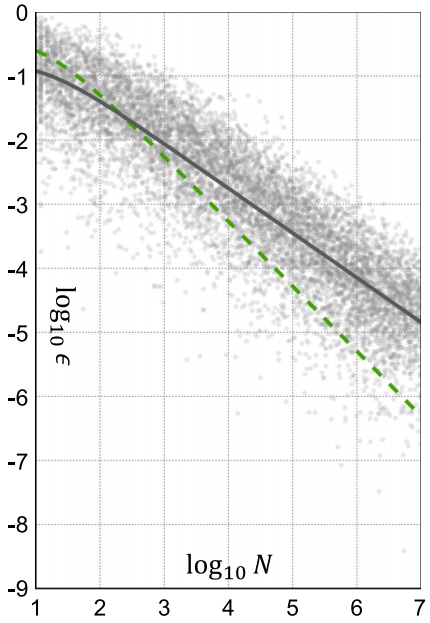

FIG. 8. Graph of N versus ε on the $\log_{10}$-$\log_{10}$ scale for EML results. The data are fitted to $\overline{\varepsilon} \simeq O(N^{-\gamma})$ with $\gamma \simeq 0.70$ (gray line). For comparison, we also draw the line (dashed green) of SSML. It is directly observed that the SSML is superior to EML.

until $M_S^{(n)}$ becomes equal to M_E. The EML simulations are performed by varying the parameters α and β. In Fig. 8, we present the best results. Here, we get $\overline{\varepsilon} \simeq O(N^{-0.7})$. Clearly, the result is inferior to those of the SSML method. The reason is as follows. Even when $F \simeq 1$, one can arrive at the situation $M_S^{(n)} > M_S^{(n+1)}$, because the parameter update, i.e., learning, is performed by the random vector $\mathbf{r}$. Thus the resources of the state copy do not need to be wasted when F is not close to 1. Noting the aforementioned, one can infer that it is an optimal (i.e., resource efficient) strategy to perform the single-shot measurements until the failure event appears.

[1] M. G. A. Paris and J. Řeháček, *Quantum State Estimation*, Lecture Notes in Physics Vol. 649 (Springer, Heidelberg, 2004).

[2] Y. S. Teo, *Introduction to Quantum-State Estimation* (World Scientific Publishing Co., Singapore, 2015).

[3] U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, Cambridge, UK, 1997).

[4] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).

[5] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro, Phys. Rev. A **66**, 012303 (2002).

[6] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodriguez, Phys. Rev. A **69**, 010304 (2004).

[7] E. Bagan, M. A. Ballester, R. D. Gill, R. Muñoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **97**, 130501 (2006).

[8] A. Hayashi, T. Hashimoto, and M. Horibe, Phys. Rev. A **72**, 032325 (2005).

[9] J. Bae and A. Acín, Phys. Rev. Lett. **97**, 030402 (2006).

[10] C. J. Happ and M. Freyberger, Phys. Rev. A **78**, 064303 (2008).

[11] T. Sugiyama, P. S. Turner, and M. Murao, Phys. Rev. A **85**, 052107 (2012).

[12] F. Huszár and N. M. T. Houlsby, Phys. Rev. A **85**, 052120 (2012).

[13] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, Phys. Rev. A **93**, 012103 (2016).

[14] Q. Yin, L. Li, X. Xiang, G. Y. Xiang, C. F. Li, and G. C. Guo, Phys. Rev. A **95**, 012129 (2017).

[15] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, Phys. Rev. Lett. **111**, 183601 (2013).

[16] T. Opatrný, D. G. Welsch, and W. Vogel, Phys. Rev. A **56**, 1788 (1997).

[17] Y. S. Teo, H. Zhu, B. G. Englert, J. Řeháček, and Z. Hradil, Phys. Rev. Lett. **107**, 020404 (2011).

[18] K. S. Kravtsov, S. S. Straupe, I. V. Radchenko, N. M. T. Houlsby, F. Huszár, and S. P. Kulik, Phys. Rev. A **87**, 062122 (2013).

[19] R. J. Chapman, C. Ferrie, and A. Peruzzo, Phys. Rev. Lett. **117**, 040402 (2016).

[20] C. Ferrie, Phys. Rev. Lett. **113**, 190404 (2014).

[21] J. Bang, J. Lim, M. S. Kim, and J. Lee, arXiv:0803.2976.

[22] J. Bang, S. W. Lee, and H. Jeong, Quantum Inf. Process. **14**, 3933 (2015).

[23] J. Bang, J. Ryu, S. Yoo, M. Pawłowski, and J. Lee, New J. Phys. **16**, 073017 (2014).

[24] A. Hentschel and B. C. Sanders, Phys. Rev. Lett. **104**, 063603 (2010).

[25] This algorithm was originally introduced in Refs. [21,22], yet here the algorithm has been more improved with slightly changed rule for memory.

[26] Actually, the SSML shows better performance compared to the corresponding ensemble learning strategy (see Appendix C).

[27] P. Langley, *Elements of Machine Learning* (Morgan Kaufmann, San Francisco, CA, 1995).

[28] Z. B. Zabinsky and R. L. Smith, Math. Program. **53**, 323 (1992).

[29] P. Palittapongarnpim, P. Wittek, and B. C. Sanders, arXiv:1608.06238.

[30] Z. Hou, H. Zhu, G. Y. Xiang, C. F. Li, and G. C. Guo, npj Quantum Inf. **2**, 16001 (2016).

[31] B. Qi, Z. Hou, Y. Wang, D. Dong, H. S. Zhong, L. Li, and G. C. Guo, npj Quantum Inf. **3**, 19 (2017).

[32] Here, C is a constant factor and N_0 is an offset of the initial resource requirement

[33] D. Oren, M. Mutzafi, Y. C. Eldar, and M. Segev, Optica **4**, 993 (2017).

[34] Thus, for the worst case—i.e., there is no Bloch vector direction for the random mixture $\hat{\mathbb{1}}/2$, it is impractical to estimate the state.

[35] H. S. Park and A. Sharma, Opt. Fib. Tech. **32**, 102 (2016).

[36] Y. C. Liang, D. Kaszlikowski, B. G. Englert, L. C. Kwek, and C. H. Oh, Phys. Rev. A **68**, 022324 (2003).

[37] Y. I. Bogdanov, S. P. Kulik, E. V. E. Moreva, I. V. Tikhonov, and A. K. Gavrichenko, JETP Lett. **91**, 686 (2010).

[38] J. N. Damask, *Polarization Optics in Telecommunications*, Vol. 101 (Springer Science and Business Media, New York, 2004).

[39] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, Adv. At., Mol., Opt. Phys. **52**, 105 (2005).

[40] This assumption is nontrivial as the parameter space is completely characterized by the chosen set of generators, i.e., $\mathbf{G} = (\hat{G}_1, \hat{G}_1, \ldots, \hat{G}_{d^2-1})^T$, defined in Eq. (6) of the main manuscript. However, note that in the case where $d = 2$ [i.e., Bloch space spanned by Pauli SU(2) generators], the parameter space is isotropic.