Adaptive compressive tomography: A numerical study

D. Ahn, Y. S. Teo,* and H. Jeong
Department of Physics and Astronomy, Seoul National University, 08826 Seoul, Korea

D. Koutný, J. Řeháček, and Z. Hradil
Department of Optics, Palacký University, 17 listopadu 12, 77146 Olomouc, Czech Republic

G. Leuchs and L. L. Sánchez-Soto
Max-Planck-Institut für die Physik des Lichts, Staudtstraße 2, 91058 Erlangen, Germany

(Received 4 May 2019; published 26 July 2019)

We perform several numerical studies for our recently published adaptive compressive tomography scheme [D. Ahn *et al.*, Phys. Rev. Lett. **122**, 100404 (2019)], which significantly reduces the number of measurement settings to unambiguously reconstruct any rank-deficient state without any *a priori* knowledge besides its dimension. We show that both entangled and product bases chosen by our adaptive scheme perform comparably well with recently known compressed-sensing element-probing measurements, and also beat random measurement bases for low-rank quantum states. We also numerically conjecture asymptotic scaling behaviors for this number as a function of the state rank for our adaptive schemes. These scaling formulas appear to be independent of the Hilbert-space dimension. As a natural development, we establish a faster hybrid compressive scheme that first chooses random bases, and later adaptive bases as the scheme progresses. As an epilogue, we reiterate important elements of informational completeness for our adaptive scheme.

DOI: 10.1103/PhysRevA.100.012346

I. INTRODUCTION

The main aim of quantum-state tomography [1–3] is to reconstruct the unknown true quantum state $\rho_{\rm t}$ from data $\mathbb{D}$ obtained after a set of measurements (collectively represented by the map $\mathscr{M}[\rho_{\rm t}] \mapsto \mathbb{D}$) are performed. Since quantum mechanics constrains $\rho_{\rm t}$ to be a positive operator of unit trace, it is unambiguously specified by $d^2 - 1$ free parameters.

To fully characterize $\rho_{\rm t}$, one performs an informationally complete (IC) set of $O(d^2)$ measurement outcomes $\Pi_j \geqslant 0$ that form a positive operator-valued measure (POVM) ($\sum_j \Pi_j = 1$). The corresponding physical probabilities $p_{{\rm t},j}$ relate the POVM and $\rho_{\rm t}$ via Born's Rule, expressed by either $p_{{\rm t},j} = {\rm tr}\{\rho_{\rm t}\,\Pi_j\}$ or $\mathscr{M}[\rho_{\rm t}] = \boldsymbol{p}_{\rm t}$. Realistically, $\boldsymbol{p}_{\rm t}$ is always inaccessible. Rather, $\mathbb{D}$ is collected from $N < \infty$ independent sampling copies, so that one acquires the relative frequency $\nu_j = n_j/N$ for each outcome Π_j of n_j data counts. Working with $\mathbb{D} \equiv \{\nu_j\}$, the *maximum-likelihood* (ML) method [3–6] may be used to supply the unique state estimator $\widehat{\rho}_{\rm ML} \geqslant 0$ and its corresponding ML physical probabilities ($\boldsymbol{p}_{\rm ML}$). In the limit $N \gg 1$, we have $\widehat{\rho}_{\rm ML} \to \rho_{\rm t}$. In both hypothetical noiseless and practical noisy situations, we say that operationally, $\mathbb{D}$ *is IC if* $\widehat{\rho}_{\rm ML}$ *is unique with respect to the whole state space* [7,8]. Incidentally, this notion is synonymous to that of "strictly complete" in [9,10].

For complex quantum systems in high-dimensional states, measuring $O(d^2)$ outcomes quickly turns into a resource-intensive practical problem [11,12]. A prototypical strategy to circumvent this problem is to first assume *a priori* that ${\rm rank}\{\rho_{\rm t}\} \leqslant r$ for a given small $r \ll d$, and next invoke the conventional method of *compressed sensing* (CS) [13–15] to find a unique rank-deficient state estimator [9,10,16–18]. This widely held standard, however, faces two practical issues. To start off, the *a priori* assumption about r requires additional verification before it can be applied to CS since the resulting estimator accuracy hinges on the validity of this assumption. Next, as one has no means of validating the final estimator self-consistently, the usual solution is to compare the estimator with some assumed target state [9,17,18]. In the presence of experimental errors, there is simply no guarantee whether such a comparison is actually trustworthy.

In Ref. [19], we developed and experimentally carried out an adaptive compressive tomography (ACT) protocol that does not rely on any *a priori* information about the quantum state apart from its dimension d. It consists of a self-consistent informational completeness certification (ICC) stage and an adaptive measurement optimization stage, both of which operate only on accumulated data. For practical implementation of ACT, we considered adaptive choices of von Neumann (orthonormal) measurement bases that are generally more easily realized in the laboratory than arbitrary POVMs, as well as the product ACT (PACT) that is compatible with many-body systems by utilizing local product bases instead of entangled ones as in ACT.

In this article, we shall perform a series of numerical studies to explore (P)ACT. In particular, (a) we shall demonstrate, up to a reasonably large d and r, that in terms of the number of IC bases $k_{\rm IC}$ needed for unique reconstructions, our adaptive schemes outperform known random bases measurements.

*ys_teo@snu.ac.kr

(b) We next show that the $k_{\rm IC}$ scaling behaviors of both ACT and PACT match the respective ones of the element-probing Baldwin-Flammia POVMs and Baldwin-Goyeneche bases reported in [10], and thereby effectively provides conjectured asymptotic scaling behaviors for both adaptive schemes in the limits $r \ll d$ and $d \gg 2$. More specifically, we have $k_{\rm IC} = 2r + 2$ for ACT and $k_{\rm IC} = 4r + 1$ for PACT. Finally, (c) we present a "random-adaptive" hybrid compressive tomography scheme that combines the respective speed and compressive advantages of random and adaptive strategies.

The article is organized as follows. Sections II and III, respectively, provide a preliminary introduction to CS and ACT for subsequent discussions. After providing some details about the simulation specifications in Sec. IV, all numerical results are then presented in Sec. V. We end the article with additional remarks on the concept of "IC" in ACT in Sec. VI.

II. COMPRESSED-SENSING TOMOGRAPHY

The concept of CS [13–15] provides one approach to recover an unknown signal *of known degree of sparsity* by performing a small set of specialized compressive measurement outcomes, and has since been widely adopted in signal processing [20]. In the context of quantum-state tomography [9,10,16–18], the low-rank assumption often taken for granted, that is, $\mathrm{rank}\{\rho_{\rm t}\} \leqslant r$ for a *known* and sufficiently small integer r, permits the utilization of CS to uniquely reconstruct $\rho_{\rm t}$ with very few outcomes. While the primary focus began with random Pauli observables $M = O[rd(\ln d)^2]$ [16,21], compressive measurements of more variety were later constructed [10].

A. Summary of the compressed-sensing procedure

Based on the assumed prior information, for measurement map $\mathscr{M}$ to be compressive and IC with respect to $\rho_{\rm t}$, it is sufficient to satisfy the so-called restricted isometry property [22,23]. Without spelling out unnecessary mathematical details, we note that there also exist other kinds of compressive measurements without such a property [10]. With a compressive $\mathscr{M}$, one seeks the unique estimator $\widehat{\rho}$ from the data convex set $\mathscr{C}$, which contains all possible states that satisfy the constraint $\mathscr{M}[\widehat{\rho}] = \boldsymbol{p}_{\rm t}$ in the noiseless regime.

One possible way is to choose $\widehat{\rho}$ that has the lowest rank out of $\mathscr{C}$. It can be shown that for $\mathscr{M}$ satisfying the restricted isometry property above a certain threshold degree of orthonormality, this nonconvex optimization is equivalent to the convex optimization of the trace-norm (nuclear-norm) minimization [24], where the rank and the trace norm are, respectively, the operator version of the l_0 norm and the l_1 norm meant for signal recovery [14]. This ensures that the set of states consistent with $\mathbb{D}$ intersects the state space at exactly one lowest-rank state that coincides with $\rho_{\rm t}$ if $N \gg 1$. With real dataset $\mathbb{D}$ and positivity (see, also, Fig. 1), the CS algorithm (proven to work with measurements possessing restricted isometry property) that uniquely identifies a state estimator may be simplified [9] into the following:

CS procedure with positivity.
For known r and data $\mathbb{D} = \{\nu_j\} \equiv \boldsymbol{\nu}$:
(1) Minimize $\mathrm{tr}\{\rho'\}$, subject to
  (a) $\|\mathscr{M}[\rho'] - \boldsymbol{\nu}\| \leqslant \varepsilon$ for some $\varepsilon > 0$ that depends on noise,
  (b) $\rho' \geqslant 0$.
(2) Trace renormalize the optimal ρ'.

It can be shown that the above optimization routine leads to perfect recovery of $\rho_{\rm t}$ for noiseless data ($\boldsymbol{\nu} \to \boldsymbol{p}_{\rm t}$ with $\varepsilon \to 0$).

B. Typical compressive measurements

For subsequent numerical comparisons, we shall consider a few well-known compressive POVMs that were applied to quantum-state tomography. The first classic compressive measurement is the set of random Pauli (RP) bases for n-qubit systems [9,16], which maximally comprise 3^n sets of n tensor products of the standard qubit Pauli operators σ_x, σ_y, and σ_z.

On the other hand, the independent studies of two other types of measurements [25,26] related to pure-state distinction eventually led to the construction of rank-r generalized compressive measurements [10]. These are the element-probing Baldwin-Flammia (BF) POVM and Baldwin-Goyeneche (BG) bases that are constructed using the mathematical concepts of Schur complement and block diagonalization of the density matrix. More importantly, these compressive measurements have analytical performance scaling behaviors in the regime $r \ll d$. Upon denoting the number of IC outcomes as $M_{\rm IC}$, the BF POVM gives $M_{\rm IC} = (2d - r)r + 1$ [two outcomes more than the number of free parameters $(2d - r)r - 1$ for a rank-r state], and the BG bases give $k_{\rm IC} = 4r + 1$, or $M_{\rm IC} = (4r + 1)d$.

Random measurements first fueled the progress of CS [14]. For state tomography, we look into two classes of random

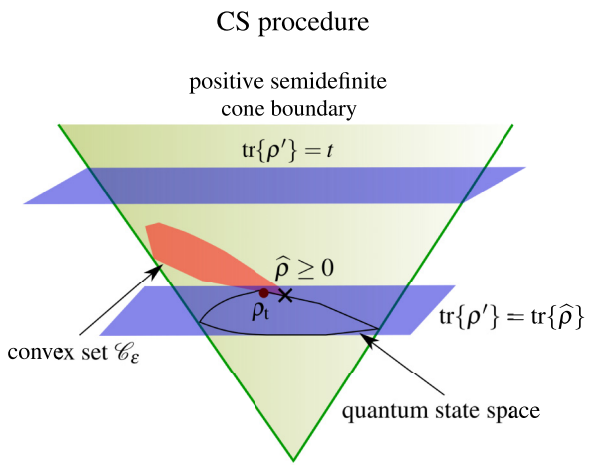

FIG. 1. Pictorial review of CS. A convex set $\mathscr{C}_\varepsilon$ is formed from quantum states that satisfy the inequality $\|\mathscr{M}[\rho'] - \boldsymbol{\nu}\| \leqslant \varepsilon$. For an appropriate compressive measurement map $\mathscr{M}$, trace minimization over the state space yields a unique state estimator $\widehat{\rho} \geqslant 0$ that is typically near $\rho_{\rm t}$ for $N \gg 1$.

measurements. The first class is the set of random bases generated by Haar unitary operators (RH) that are widely used in quantum information theory [27–30]. The second class is the set of eigenbases of random full-rank states (RS) distributed uniformly according to the Hilbert-Schmidt measure [31]. For more details regarding the numerical constructions of random bases, we refer the interested reader to the Appendix. Finally, there exists yet another benchmark by Kech and Wolf (KW) for arbitrary von Neumann bases, which states that $k \geqslant 4r\lceil\frac{d-r}{d-1}\rceil$ is sufficient to uniquely reconstruct any rank-r state [32].

C. Issues in conventional compressed sensing

While the CS procedure is a promising candidate for low-rank state recovery, there are two primary concerns that need to be addressed.

First, the proper set of compressive measurements is based on the premise that the upper bound r of the rank of ρ_t is known. The assumption of r must thus be experimentally justified. A categorical misclassification of ρ_t would result in either a non-IC (too small an r) or an unnecessarily overcomplete dataset (too big an r). The former leads to an ambiguous set of estimators, while the latter overuses measurement resources.

Second, there exists no method in CS to systematically validate if the acquired $\mathbb{D}$ is truly IC in real experiments. Instead, the fidelity measure is commonly used as the indicator that the estimator is "close" to some prechosen target state. This approach, at the very least, requires yet another round of certification for these target states and is evidently not the correct informational completeness indicator as it registers no information about the data convex set $\mathscr{C}$.

III. ADAPTIVE COMPRESSIVE TOMOGRAPHY

Our recently proposed ACT [19] is able to deterministically compress IC datasets for any given rank-r $\rho_t \equiv \rho_r$ without relying on any *a priori* information or *ad hoc* assumptions about ρ_r apart from its dimension d. The procedure of ACT is iterative and involves two main stages in every step. The first *informational completeness certification* (ICC) stage decides, given the measured outcomes and accumulated data $\mathbb{D}$, whether or not the data convex set $\mathscr{C}$, which is again the set of states $\mathscr{C} = \{\rho'\}$ for which $\mathscr{M}[\rho'] = \boldsymbol{p}_{\mathrm{ML}}$, is singleton. The second *adaptive strategy* provides protocols which adaptively choose the next measurement at each step of ACT according to the accumulated data $\mathbb{D}$. The former would imply a unique state estimator consistent with an IC $\mathbb{D}$, and the procedure terminates. The latter would introduce appropriate protocols which significantly reduces the size of IC $\mathbb{D}$ compared to $O(d^2)$.

Throughout the article, we shall take the POVM to be sets of von Neumann bases that are each denoted by $\mathscr{B}$ and contains d orthonormal projectors, which are the typical kind of measurement employed in laboratory. A particular iterative adaptation step of ACT is then to search for an optimal basis to measure in the next step (see Fig. 2).

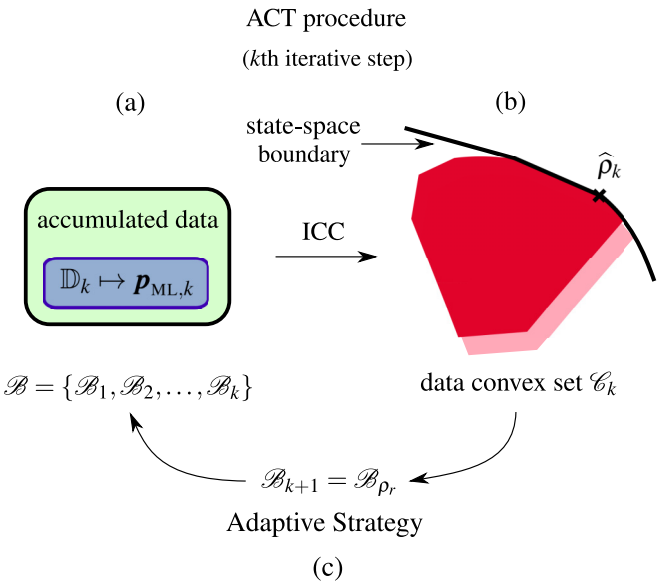

FIG. 2. Pictorial review of ACT at the kth iterative step. (a) The accumulated dataset $\mathbb{D}_k$ thus far undergoes ICC to verify whether or not (b) the data convex set $\mathscr{C}_k$ is a point. If not, (c) ACT undergoes an adaptive strategy that picks an optimal estimator $\widehat{\rho}_k$ and assigns its eigenbasis as the next von Neumann basis $\mathscr{B}_{k+1}$ to be measured. The cycle continues until $\mathscr{C}_{k=k_{\mathrm{IC}}}$ becomes a point.

A. Informational completeness certification

As ACT progresses iteratively and a sequence of von Neumann bases $\mathscr{B} = \{\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_k\}$ is measured, the size ζ_k of the data convex set $\mathscr{C}_k$ at the kth iterative step of ACT indicates whether or not the corresponding dataset $\mathbb{D}_k$ of k measured orthonormal bases is IC—$\mathbb{D}_k$ is IC if and only if $\zeta_k = 0$. In this case, we denote the set of IC bases to be $k_{\mathrm{IC}} = k$. In other words, the task of ICC is to verify if $\zeta = 0$ (or close to some small numerical threshold value in practice).

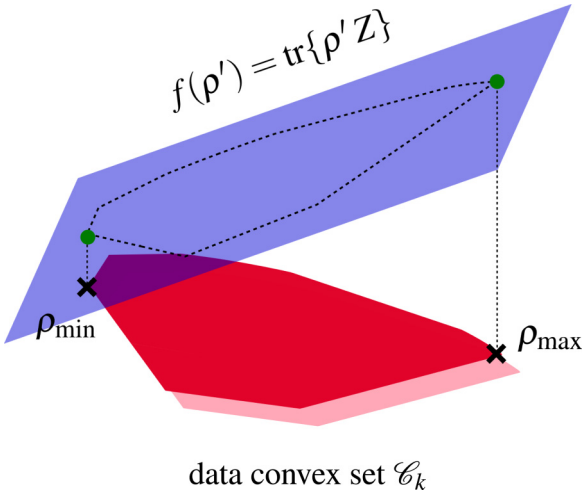

FIG. 3. The ICC procedure. The linear function $f(\rho') = \mathrm{tr}\{\rho' Z\}$ (represented by a hyperplane) has global optimas at the edges of $\mathscr{C}_k$. The corresponding auxiliary extrema states ρ_{max} and ρ_{min} determine whether or not $\mathscr{C}_k$ is a singleton through $s_{\mathrm{CVX},k}$.

Fortunately, while ζ_k is a complicated state-space integral with respect to some volume measure over $\mathscr{C}_k$, there is an alternatively much more feasible way to detect if $\zeta_k = 0$. Using a randomly chosen full-rank state Z throughout the entire run of ACT, if we define the linear function $f(\rho') = \mathrm{tr}\{\rho' Z\}$ for the variable state ρ', then, owing to the convexity of $\mathscr{C}_k$ *regardless of whether data noise is present*, it straightforwardly follows that $\zeta_k = 0$ iff $f_{\min,k} = f_{\max,k}$, where $f_{\min,k}$ and $f_{\max,k}$ are, respectively, the minimum and maximum values of f in $\mathscr{C}_k$ (see Fig. 3). Hence, the validity of this if-and-only-if statement is robust against noise. This simple conclusion holds as long as Z is neither $1/d$ nor possesses a kernel that contains $\rho' \in \mathscr{C}_k$, which are all measure-zero situations. Simply put, the business of ICC is a semidefinite program [33] that evaluates the quantity $s_{\mathrm{CVX},k} = (f_{\max,k} - f_{\min,k})/(f_{\max,1} - f_{\min,1})$ at every kth iterative step of ACT.

ICC in the *k*th step.

(1) Maximize and minimize $f_Z(\rho') = \mathrm{tr}\{\rho' Z\}$ for a fixed, randomly chosen full-rank state $Z \neq 1/d$ to obtain $f_{\max,k}$ and $f_{\min,k}$ subject to
   (a) $\rho' \geqslant 0$, $\mathrm{tr}\{\rho'\} = 1$,
   (b) $\mathscr{M}[\rho'] = \boldsymbol{p}_{\mathrm{ML},k}$ for $\boldsymbol{p}_{\mathrm{ML},k}$ obtained from the accumulated $\mathbb{D}_k$ thus far.

(2) Compute $0 \leqslant s_{\mathrm{CVX},k} \leqslant 1$ and check if it is smaller than some threshold ε.

(3) If $s_{\mathrm{CVX},k} < \varepsilon$, terminate ACT. Continue otherwise.

More specifically, s_{CVX} is known as the *size monotone* in the sense that with decreasing ζ_k or increasing k, $s_{\mathrm{CVX},k}$ never increases in value for near-perfect $\mathbb{D}_k$. The reason for this is that with perfect $\mathbb{D}$s, we have the set hierarchy $\mathscr{C}_1 \supseteq \mathscr{C}_2 \supseteq \cdots \supseteq \mathscr{C}_{k_{\mathrm{IC}}}$ as more linearly independent bases are measured, so that $f_{\max,k}$ monotonically decreases and $f_{\min,k}$ monotonically increases for a linear f, or $s_{\mathrm{CVX},k+1} \leqslant s_{\mathrm{CVX},k}$. Moreover, it is easy to see that if $f_{\max,k+1} - f_{\min,k+1} < f_{\max,k} - f_{\min,k}$, then $\mathscr{C}_{k+1} \subset \mathscr{C}_k$.

In terms of complexity, unlike the traditional NP-hard problem to determine whether a given set of POVM possesses the restricted isometry property [34], the complexity to decide if this POVM is IC with respect to $\mathbb{D}$ turns out to be, understandably, only as high as that for the semidefinite program employed in ICC.

B. Adaptive strategy

Given the collected dataset $\mathbb{D}_k$ and corresponding nonsingleton data convex set $\mathscr{C}_k$ at the kth iterative step of ACT, one seeks the optimal basis to be measured in the $(k+1)$th step in order to converge ACT at a reasonably quick pace to the singleton $\mathscr{C}_{k=k_{\mathrm{IC}}}$.

If ρ_r is a low-rank state ($r \ll d$), the natural motivation for a compressive adaptive strategy, given no other assumption, would be to actively seek the lowest-rank estimator $\widehat{\rho}_k$ from the data convex set $\mathscr{C}_k$ that ultimately either converges to ρ_t for noiseless data or is very close to it for $N \gg 1$. This is akin to conventional CS where, for a properly chosen compressive measurement map $\mathscr{M}$, the unique CS estimator, which converges to ρ_r for noiseless data, essentially has the lowest rank out of the data convex set. In order to establish a feasible scheme for finding the lowest-rank $\widehat{\rho}_k$ in $\mathscr{C}_k$, we minimize a concave function over $\mathscr{C}_k$ that has rank-minimizing characteristics. For this purpose, two exemplifying concave functions with a global minimum for pure states shall be numerically studied. They are the von Neumann entropy function $S(\rho') = -\mathrm{tr}\{\rho' \ln \rho'\}$ and the linear entropy function $S_{\mathrm{L}}(\rho') = 1 - \mathrm{tr}\{\rho'^2\}$, both of which are similar in value near all pure states and take the minimum value 0 for the pure states. We mention that entropy minimization was numerically reported in [35] to offer a stronger compressive recovery for general low-rank matrices compared to the standard trace-norm minimization in CS.

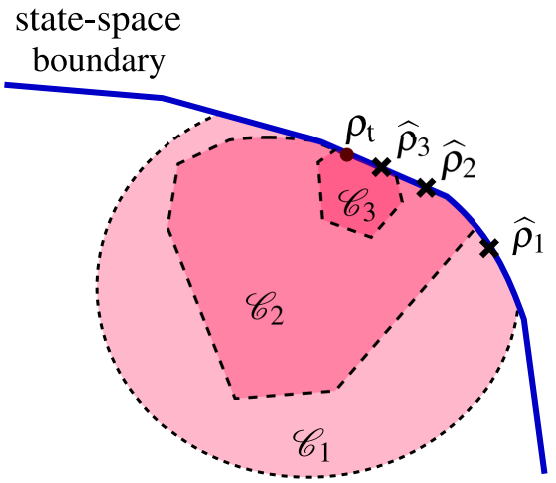

FIG. 4. The progress of the adaptive strategy in ACT. In each step, (linear-)entropy minimization is performed under both positivity and data (ML) constraints to give lowest-(linear-)entropy $\widehat{\rho}_k$. The data convex set $\mathscr{C}_k$ makes a common boundary with the state space. This boundary shrinks as k increases, so that the $\widehat{\rho}_k$'s converge to ρ_r.

While in CS the measurement map is first decided before a unique lowest-rank estimator is obtained, we note here that for ACT, the mechanism that drives the compressive nature is the interplay between positivity and data (ML) constraints on the entropy-minimization procedure, which leads to the correct guidance to the unique $\widehat{\rho}_{k_{\mathrm{IC}}} \approx \rho_r$. It is shown in [19] that owing to these two constraints, if one manages to measure the eigenbasis $\mathscr{B}_{\rho_r}$ of ρ_r that is low rank ($r \ll d$), then only $k_0 = \lceil (r^2 - r)/(d-1) \rceil + 1 \ll d + 1$ bases (including $\mathscr{B}_{\rho_r}$) are needed to unambiguously reconstruct ρ_r. Therefore, the eigenbasis of a lowest-(linear-)entropy $\widehat{\rho}_k$ (which would essentially have an extremely low rank) in $\mathscr{C}_k$ may be assigned to $\mathscr{B}_{k+1}$. In this manner, if $N \gg 1$, it is evident that $\mathscr{B}_{k+1}$ approaches $\mathscr{B}_{\rho_r}$, and the action of the optimization procedure under both constraints aids in speeding up this convergence (see Fig. 4).

Generally, the adaptive strategy discussed here yields a sequence of entangled bases. For many-body quantum systems, such bases are often difficult to be realized experimentally. To carry out the adaptive strategy for such systems, we impose, in every iterative step, an additional local structure on the adaptive von Neumann bases by searching for the nearest tensor-product basis that is "closest" to the optimal entangled one. One way to do so is to first minimize the distance between

the lowest-(linear-)entropy $\widehat{\rho}_k \in \mathscr{C}_k$ and another state ρ' with product eigenbasis with respect to some metric, and take the eigenbasis of the optimum. This distance reduction helps to find a local von Neumann basis that is close to $\mathscr{B}_{\widehat{\rho}_k}$ for the lowest-(linear-)entropy $\widehat{\rho}_k$.

Putting everything together, we arrive at the complete (P)ACT procedure.

(P)ACT procedure.
Beginning with $k = 1$ and a random computational basis $\mathscr{B}_1$:

(1) Measure $\mathscr{B}_k$ and collect the relative frequency data $\sum_{j'=0}^{d-1} \nu_{j'k} = 1$.
(2) From $\{\nu_{0k'}, \ldots, \nu_{d-1\,k'}\}_{k'=1}^{k}$, obtain kd physical ML probabilities.
(3) Perform ICC with the ML probabilities and compute $s_{\mathrm{CVX},k}$:
  (a) If $s_{\mathrm{CVX},k} < \varepsilon$, terminate ACT and take $\rho_{\max} \approx \rho_{\min}$ as the estimator and report $s_{\mathrm{CVX},k}$.
  (b) Else proceed.
(4) Choose a lowest-(linear-)entropy $\widehat{\rho}_k \in \mathscr{C}_k$ in $\mathscr{C}_k$.
(5) Define $\mathscr{B}_{k+1}$ to be the eigenbasis of $\widehat{\rho}_k$ for ACT, or a local basis close to it for PACT via some prechosen distance-minimization technique.
(6) Set $k = k + 1$ and repeat.

IV. SIMULATION SPECIFICATIONS

In this brief section, we clarify all essential technical details involved in our subsequent numerical studies. We begin by summarizing the main goals of our numerical studies.

(A) Compare (P)ACT with the following random measurement schemes (refer to Sec. II B):
  (I) Random Pauli (RP) bases.
  (II) Random Haar-uniform (RH) bases.
  (III) Eigenbases of random full-rank states (RS).
(B) Benchmark (P)ACT using known analytical scalings of
  (IV) Baldwin-Flammia (BF) POVMs,
  (V) Baldwin-Goyeneche (BG) bases,
  (VI) Kech-Wolf (KW) scaling for arbitrary bases, and conjecture asymptotic k_{IC} scalings for (P)ACT.
(C) Present and numerically benchmark a new hybrid compressive scheme (HCT).

In what follows, all simulations are performed on ensembles of random quantum states of various ranks r over which important indicators such as s_{CVX} and k_{IC} are averaged. To generate these ensembles, we follow the prescriptions in [31] and distribute the random states according to the Hilbert-Schmidt measure. For a fixed r, this is done by first generating $r \times d$ matrices A_j with entries independently and identically distributed according to the standard Gaussian distribution,

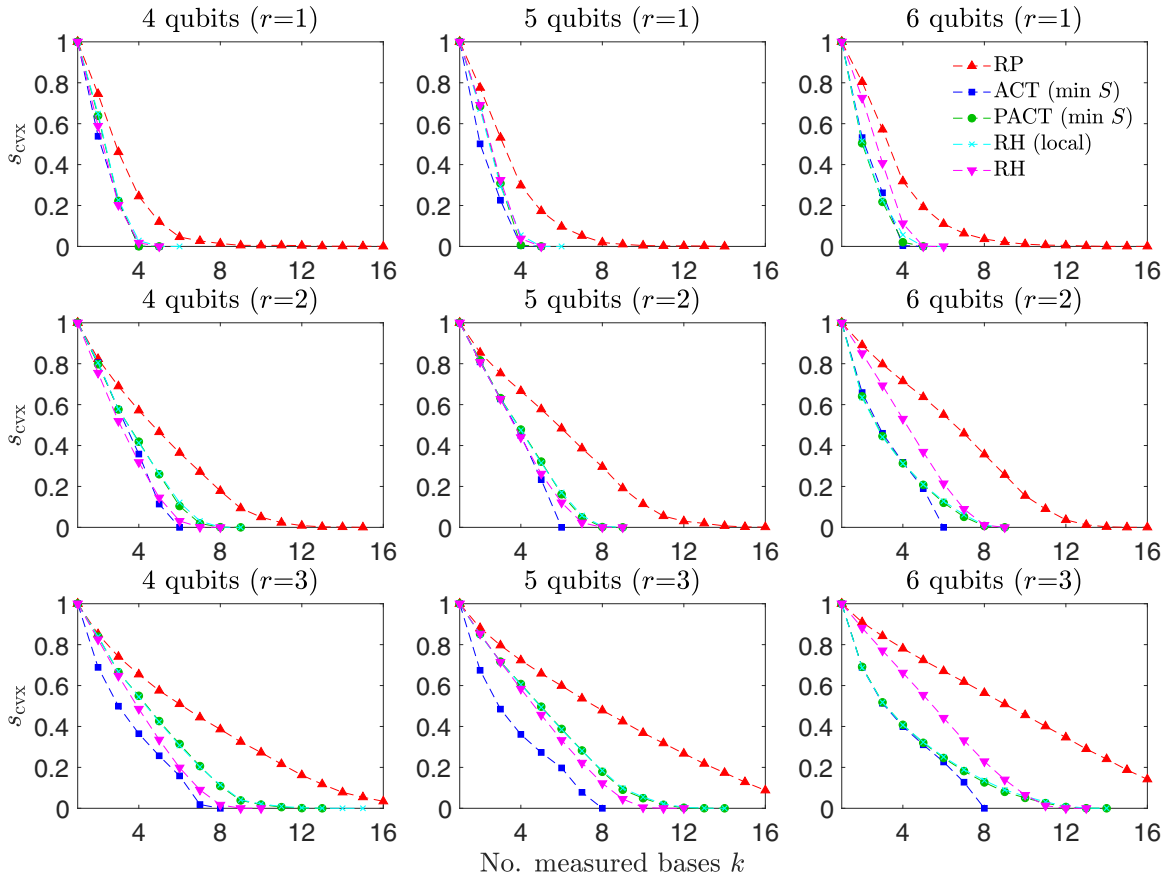

FIG. 5. Plots of s_{CVX} for multiqubit systems $d = 16$ (4 qubits), 32 (5 qubits), and 64 (6 qubits) and $1 \leqslant r \leqslant 3$. The data markers for each system and r are averaged over 100 random states that are uniformly distributed with respect to the HS measure.

and next defining the ensemble density matrices $\{\rho_j\}$ in accordance with $\rho_j = A_j^\dagger A_j / \mathrm{tr}\{A_j^\dagger A_j\}$.

Both minimum-entropy schemes (with S and S_L) over convex sets are carried out efficiently using the superfast accelerated projected gradient algorithm [6]. The ICC algorithm is carried out with the CVXOPT package [36,37]. While the deterministic compressive measurements, namely, the BF POVMs and BG bases, possess analytical k_IC scaling behaviors that can be readily used for benchmarking purposes, the random measurements, namely, RP, RH, and RS bases, have at most approximated scaling expressions. The generations of both RH and RS bases are described in the Appendix. For all *local* bases schemes, which refer to PACT, (I), (II), and (III), since the quantum systems that we shall be investigating are multiqubit, these schemes involve basis outcomes defined as projectors onto the tensor-product space of single qubits.

For the numerical simulation tasks (a)–(c), the results of which are presented in Sec. V, we study all results for noiseless data to understand the underlying theoretical relationships between $\{s_\mathrm{CVX}, k_\mathrm{IC}\}$ and r.

V. NUMERICAL RESULTS

A. Comparisons of (product) adaptive compressive tomography with random-bases measurements

We proceed to systematically compare both ACT and PACT, respectively, with popular random compressive schemes (I)–(III) in Sec. IV. Figure 5 shows the results of ICC for true quantum states of 4-, 5-, and 6-qubit systems. For each type of system, ICC is applied to all schemes on states of ranks $r = 2$, 4, and 6. For all the schemes, we confirm that $s_{\mathrm{CVX},k}$ monotonically decreases over the number of measured von Neumann bases k, and reach $s_{\mathrm{CVX},k_\mathrm{IC}} = 0$ (up to some numerical threshold). In terms of the convergence rate of $s_{\mathrm{CVX},k}$, it turns out that (P)ACT as well as schemes (II)–(V) are more efficient than RP. Here, the ACT scheme shows most efficient performance. Faster convergence of $s_{\mathrm{CVX},k}$ directly leads to smaller k_IC, which implies higher compression in the size of the IC data.

The compression efficiencies for all schemes are compared via k_IC more clearly in Fig. 6, which validates that ACT is the most efficient scheme for all numbers of qubits, whereas RP turns out to be the most inefficient one. More specifically, it is evident that the k_IC gap between ACT and RH, and that between PACT/RH (local) and RP, enlarges for larger r and number of qubits, which implies that ACT is relatively more for states of higher rank and Hilbert-space dimension. The apparent coincidence between PACT and RH (local) for all d may be attributed to an unavoidable minimum level of incurred randomness in PACT that arises from the restriction to local-bases measurement during adaptive optimization. Such a level of randomness is sufficient to practically render PACT equivalent to RH.

B. Benchmarks for (product) adaptive compressive tomography and asymptotic performances

Benchmarking of the k_IC performance for ACT, RH, RS, PACT, and local RH schemes is done with the known analytical standards (IV)–(VI) listed in Sec. IV, and the results are presented in Fig. 7. The ACT schemes carried out through minimizing S and S_L are compared with the element-probing BF POVMs for arbitrary rank-r states, which contains outcomes of nonunit rank and possesses a total number of $M = (2dr - r^2 + 1)$ outcomes. The figure shows that for both ACT protocols, k_IC scaling behaviors are in good agreement with the expression $M/d + 2$ (a shifted BF scaling). This means that ACT, which yields only rank-1 von Neumann projectors, exceeds only by two bases in performance as compared to BF POVMs to perform IC state tomography.

We also note that both ACT schemes defined by the minimization of two different entropy functions give almost indistinguishable k_IC curves for all tested r and d. This numerical observation confirms that the behaviors of von Neumann and linear entropies are almost the same with respect to the optimization algorithm in the adaptive stage. Thus, the (P)ACT protocols involving the von Neumann entropy minimization are sufficient for the remaining discussions. Indistinguishable k_IC curves also exist for random schemes (such as RH and RS), and this amply hints that the effect of informational completeness for the nonadaptive schemes depends weakly on the specific choice of the random-bases generation algorithm, but rather more strongly on the state rank.

One might, at this point, naively suppose that perhaps ACT should also work sufficiently well if $\widehat{\rho}_k$ is some randomly chosen rank-deficient state in $\mathscr{C}_k$ for every k. Figure 8 compares the average k_IC of min-S adaptive bases with that of eigenbases of random rank-deficient states. It presents clear numerical evidence contrary to the above supposition, which further reinforces the statement that unlike random schemes, an appropriate choice of objective function is necessary for ACT to achieve optimal compression.

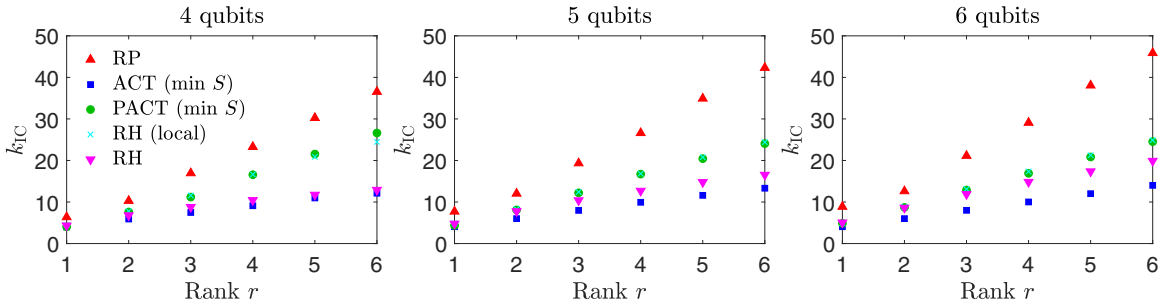

FIG. 6. Plots of simulation of k_IC, which show an aspect that is different from Fig. 5 of the ICC performance for all schemes.

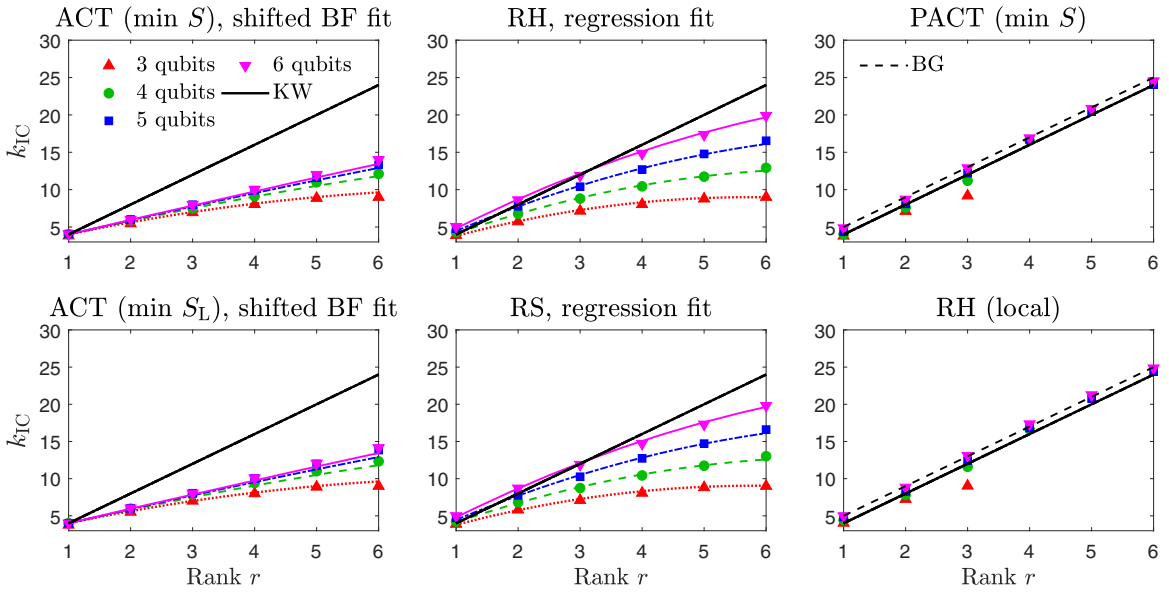

FIG. 7. A comparison of various schemes for different multiqubit systems and some of their behaviors relative to BF POVMs and BG bases. Both ACT and PACT follow closely to the respective scaling behaviors of the (shifted) BF POVMs and BG bases. The nonlinear regression fits for RH and RS follow a general $O(r \ln d)$ slope with quadratic r-dependent intercepts that are relatively weak in the limit $r \ll d$. The linear KW benchmark overestimates both ACT schemes for $r > 1$, and does so as well for RH and RS schemes. Its linear behavior lies close to that of the BG bases scheme, with the latter having very similar scaling as the asymptotic behaviors for both the PACT and RH bases scheme.

In the context of many-body local compressive measurements, we compare the k_{IC}'s of PACT and local RH scheme with those of the BG bases for arbitrary r and d, the latter of which employs $4r + 1$ specifically constructed orthonormal entangled bases. Figure 7 tells us that both product schemes asymptotically approach the BG scheme in k_{IC} performance as the number of qubits grows.

Hence, we numerically confirm that for all r and d, ACT shows stronger compressive behavior than PACT and nonadaptive random schemes. The result of benchmarking with BF POVMs and BG bases reveals k_{IC} behaviors for ACT and PACT that lose the dependency on d in the limit of a large number of qubits. These conjectured asymptotic scalings are, respectively, $k_{IC} = 2r + 2$ and $k_{IC} = 4r + 1$, and our belief in their validity is further strengthened with 7-qubit systems in Fig. 9.

All of the compressive bases schemes presented here are contrasted with the KW benchmark that applies to arbitrary bases measurements, which is effectively a linear function of r that is independent of d due to the ceiling function. We point out that the KW benchmark that estimates the required number of bases needed to perform IC reconstruction using arbitrary von Neumann bases almost always overestimates the k_{IC} performance of ACT. The reason is that the measurement resources for ACT scale nonlinearly with r owing to the low-rank guidance from both positivity and data constraints. The KW benchmark also show signs of overestimation for the random RH and RS schemes for sufficiently large r.

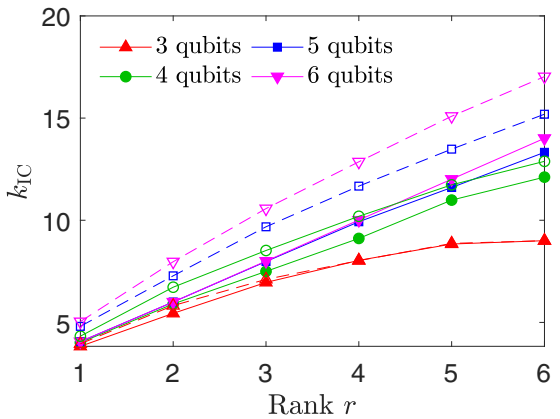

FIG. 8. Plots of k_{IC} for min-S ACT (filled markers and solid curves) and the strategy of picking eigenbases of random rank-deficient states from data convex sets (unfilled markers and dashed curves). The discrepancies are magnified for large d.

C. Hybrid compressive tomography

While the complete ACT scheme is highly compressive ($k_{IC} = 2r + 2$), the computational resources needed to search for the optimal estimators in the data convex set will

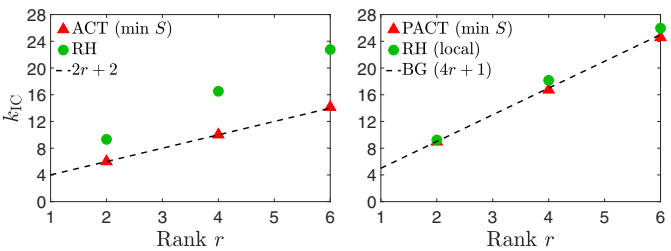

FIG. 9. Plots of k_{IC} scaling behaviors for 7-qubit systems ($d = 128$), averaged over 50 random states per data marker. The scaling conjectures $2r + 2$ and $4r + 1$ still stand accurately for the sampled r values.

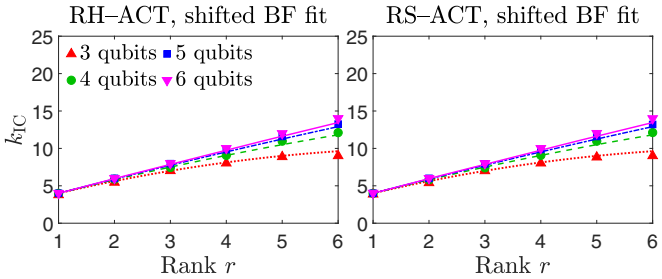

FIG. 10. Plots of k_{IC} for HCT with $s_{\mathrm{th}} = 0.5$. Both sets of graphs match almost exactly with the two ACT plots in Fig. 7.

eventually become noticeably expensive for extremely large physical systems. On the other hand, a fully random scheme simply suggests measurement bases randomly and requires essentially negligible computational resources.

As an attempt to speed up the compression process, we may take advantage of the benefits from both random and adaptive protocols and establish a hybrid compressive tomography (HCT) scheme. This hybrid scheme starts with choosing random von Neumann measurement bases, which may be constructed with either the RH or RS prescription. In fact, any sort of bases would deliver similar performance due to the limited data one always has at the beginning of a compression process. As the accumulated dataset is built up, more information is gained about the unknown ρ_r such that it is now justified to spend more effort in searching for good optimal measurements based on more statistically reliable data. Armed with this basic insight, we propose the following procedure:

HCT procedure.
Beginning with $k = 1$, a random computational basis $\mathscr{B}_1$ and some positive threshold value s_{th}:

(1) Measure $\mathscr{B}_k$ and collect the relative frequency data $\sum_{j'=0}^{d-1} \nu_{j'k} = 1$.
(2) From $\{\nu_{0k'}, \ldots, \nu_{d-1\,k'}\}_{k'=1}^{k}$, obtain kd physical ML probabilities.
(3) Perform ICC with the ML probabilities and compute $s_{\mathrm{CVX},k}$:
  (a) If $s_{\mathrm{CVX},k} < \varepsilon$, terminate ACT and take $\rho_{\max} \approx \rho_{\min}$ as the estimator and report $s_{\mathrm{CVX},k}$.
  (b) Else proceed.
(4) If $s_{\mathrm{CVX},k} > s_{\mathrm{th}}$:
  (a) Assign a random basis, perhaps from the Appendix, to $\mathscr{B}_{k+1}$.
  Else:
  (a) Assign an optimal basis obtained using the adaptive strategy in (P)ACT to $\mathscr{B}_{k+1}$.
(5) Set $k = k + 1$ and repeat.

Figure 10 reveals that both RH-ACT- and RS-ACT-type HCT schemes give almost the same averaged k_{IC} behaviors, both of which are essentially identical to ACT as in Fig. 7. Based on these numerical evidences, we may argue that the cause of these equivalent performances comes from almost identical gradual average gains in tomographic information throughout the respective compressive tomography schemes. This supports our observation that without relying on any *a priori* knowledge about ρ_r, adaptive methods only serve as crucial roles in compressive tomography after a sufficient amount of tomographic information is acquired for these methods to yield more correct optimized measurement bases. Otherwise, the identical average compression capabilities of both ACT and HCT imply that both schemes effectively acquire the same amount of tomographic information at the early stages of the processes. This justifies the more economical approach of first measuring random bases during the initial stage of HCT, which is far less time consuming than optimizing for adaptive bases in ACT on highly complex quantum systems.

VI. EPILOGUE: REMARKS ON INFORMATIONAL COMPLETENESS IN ADAPTIVE COMPRESSIVE TOMOGRAPHY

Besides s_{CVX}, we may inspect other aspects of the data convex set in ACT, carried out with S minimization for example, to preempt its correct termination in case one intends to stop the reconstruction procedure with $k < k_{\mathrm{IC}}$ measurement bases. Figure 11 showcases the average behaviors of various different minimum-entropy $\widehat{\rho}_k$ qualities for ACT under negligible statistical fluctuation ($N \gg 1$). We see that the fidelity, rank, and entropy of $\widehat{\rho}_k$ reach the correct true values as k increases, and saturate at $k < k_{\mathrm{IC}}$. Technically, both the fidelity and rank cannot be used to judge the termination of ACT since the fidelity requires knowledge of ρ_r, and it is also impossible to guess whether the rank of $\widehat{\rho}_k$ is close to the correct value without rank$\{\rho_r\}$ as a reference because of its regular stepped gradient. However, based on our numerical experience, the monotonically increasing entropy $S(\widehat{\rho}_k)$ (a direct mathematical consequence of the inequality chain $\mathscr{C}_1 \supseteq \mathscr{C}_2 \supseteq \ldots \supseteq \mathscr{C}_{k_{\mathrm{IC}}}$ for noiseless data) always approaches the final true value with smoothly decreasing gradient $S(\widehat{\rho}_k) - S(\widehat{\rho}_{k-1}) < S(\widehat{\rho}_{k-1}) - S(\widehat{\rho}_{k-2})$. We may then attempt to prematurely stop ACT at $k = k_0$ when $S(\widehat{\rho}_{k_0}) - S(\widehat{\rho}_{k_0-1})$ is small enough and use the resulting estimator $\widehat{\rho}_{k_0}$ as an *approximately IC reconstruction* for future statistical predictions, in which case this low-rank $\widehat{\rho}_{k_0}$ will be close to ρ_r even though the size of $\mathscr{C}_{k_0}$ can be appreciably greater than zero. With statistical consistency, this termination method also works for real data of sufficiently large N (or low statistical fluctuation).

The aforementioned remarks eventually lead to a subtle point behind the notion of "IC" adopted by ACT. In general tomography settings, which include the studies in CS, one typically defines "IC" in terms of the class of quantum states of interest, such as the class of rank-r states to name one. In ACT, however, we speak of an IC dataset $\mathbb{D}$ that is statistically related to the particular unknown state we are probing. As a result, the notion of IC in ACT, along with the value of k_{IC}, strongly depends on both the adaptive bases and $\mathbb{D}$.

A more objective $\mathbb{D}$- or ρ_r-dependent compressive strategy in exchange for spurious or *ad hoc* prior assumptions about ρ_r also means that an ACT bases POVM that is IC for a given unknown rank-r state is not necessarily going to be IC for another rank-r state. Figure 12 elucidates this fact for a

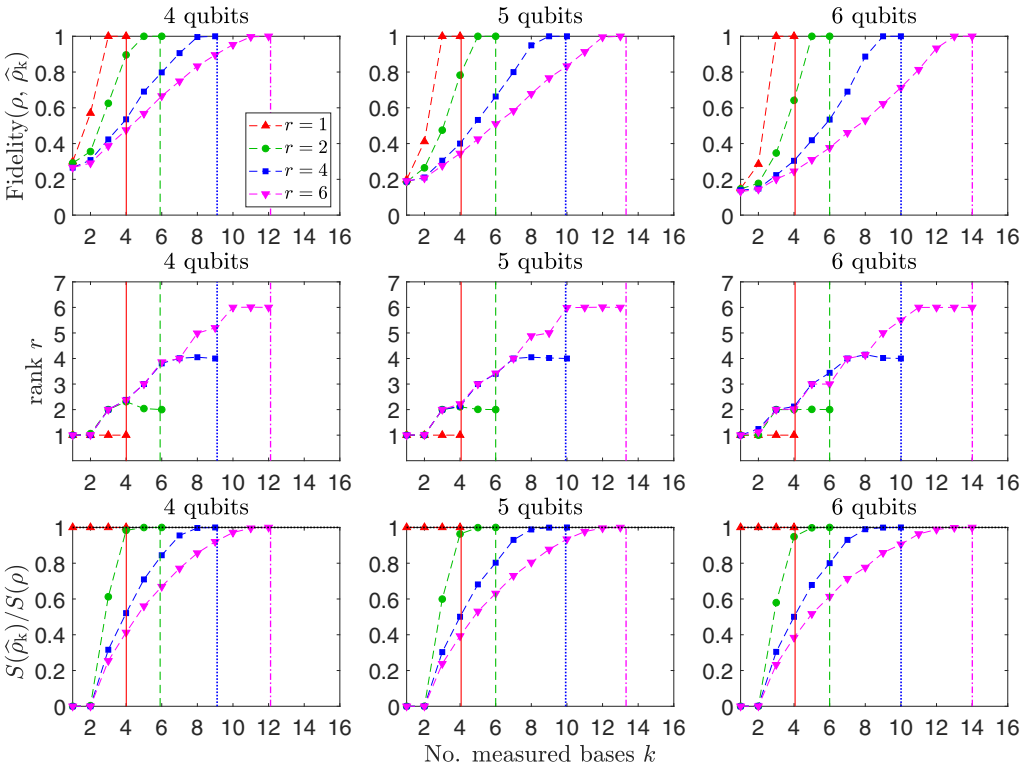

FIG. 11. Noiseless evolution of the different minimum-entropy estimator qualities as ACT progresses. Vertical lines mark the k_{IC} values for the respective ranks r. All curves are averaged over 100 random states that are uniform in the HS measure per r value. Here, $S(\rho)$ is only a convenient normalization factor for averaging $S(\widehat{\rho}_k)$ over true states.

distribution of rank-one true states probed by an ACT bases POVM that is IC for a particular pure state assuming noiseless probabilities.

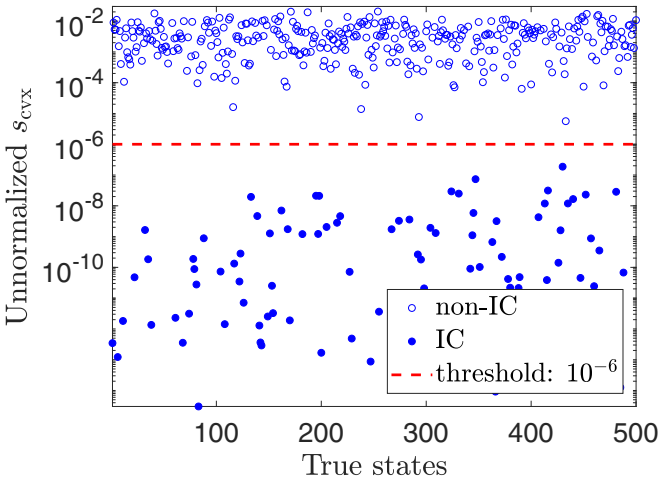

FIG. 12. Noiseless distribution of s_{CVX} for a Haar-uniform set of 500 rank-one 5-qubit ρ_r's ($d = 2^5 = 32$) based on a fixed set of ACT bases that is IC for another pure state. The decision for labeling the probabilities as "IC" typically only weakly relies on the choice of a threshold so long as it is reasonable. Since truly IC situations correspond to very small s_{CVX} close to the numerical limits of the semidefinite program in ICC, which are clearly separated from IC situations of significantly larger s_{CVX} as shown in this figure, a reasonable threshold value below which all points are IC may be conveniently fixed at 10^{-6}.

VII. CONCLUSIONS

We have performed a comprehensive numerical study of our adaptive compressive tomography schemes, which uniquely reconstruct any arbitrary rank-deficient quantum state with few measurement bases without any other *a priori* information about this state. Several numerical results in this article can serve as important guidelines for resource planning, structuring, and executing objective compressive tomography for multiqubit systems, which are widely used in quantum information and quantum optics studies.

After simulating up to reasonably large dimensions d of multiqubit systems and quantum-state rank r, we can summarize these numerical results under low statistical fluctuation:

(1) The average performances (minimal number of informationally complete bases k_{IC}) of adaptive compressive tomography schemes almost always beat those of random compressive schemes, which include random Pauli bases, random Haar-unitary bases and eigenbases of random Hilbert-Schmidt-uniform states.

(2) The average $M_{\mathrm{IC}} = k_{\mathrm{IC}}d$ for the entangled adaptive compressive scheme is greater than that of Baldwin-Flammia measurement outcomes by only $2d$ for all of the tested range of r and d. The asymptotic ($d \gg 2$) average k_{IC} for both entangled and product adaptive schemes are, respectively, $2r + 2$ and $4r + 1$.

(3) There is virtually no difference in the average k_{IC} performance between a fully adaptive scheme and a hybrid scheme that starts off as a random scheme followed by an adaptive scheme at a later stage for a reasonable

transition point. Therefore, the hybrid compressive scheme may be used to speed up the tomography compression process.

ACKNOWLEDGMENTS

We acknowledge financial support from the BK21 Plus Program (Grant No. 21A20131111123) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF), the NRF grants of Korea funded by the Ministry of Science and ICT (Grants No. NRF-2019R1H1A3079890 and No. NRF-2018K2A9A1A06069933), the European Research Council (Advanced Grant PACART), the Spanish MINECO (Grant No. FIS2015-67963-P), the Grant Agency of the Czech Republic (Grant No. 18-04291S), and the IGA Project of the Palacký University (Grant No. IGA PrF 2019-007).

APPENDIX: CONSTRUCTIONS OF RANDOM BASES

We briefly supply two short routines to construct von Neumann bases that are generated by random Haar-distributed unitary operators (RH bases), as well as those that are eigenbases of random full-rank states uniformly distributed according to the Hilbert-Schmidt measure (RS bases). These two unitary sets have generally different operator probability distributions, which may be verified by comparing some of their operator moments.

It is well known that the *QR* decomposition generates unitary operators distributed according to the Haar measure [27], so that the following routine applies:

Constructing an RH basis.
Starting from a reference basis $\mathscr{B}_{\rm ref} = \{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$:

(1) Generate a random $d \times d$ matrix A with entries an independent and identically distributed standard Gaussian distribution.
(2) Compute Q and R from the QR decomposition $A = QR$.
(3) Define $R_{\rm diag} = {\rm diag}\{R\}$.
(4) Define $L = R_{\rm diag} \oslash |R_{\rm diag}|$ ($\oslash$ refers to the Hadamard division).
(5) Define $U_{\rm Haar} = QL$.

Thereafter, construct the new basis $\mathscr{B}_{\rm Haar} = U_{\rm Haar}\mathscr{B}_{\rm ref}$.

Steps (3)–(5) enforce a *QR* decomposition procedure that produces an effective "*R*" matrix that has positive diagonal entries, which is the correct decomposition procedure we need to generate $U_{\rm Haar}$.

The second class may be easily generated by the following pseudocode:

Constructing an RS basis.
Starting from a reference basis $\mathscr{B}_{\rm ref} = \{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$:

(1) Generate a random $d \times d$ matrix A with entries an independent and identically distributed standard Gaussian distribution.
(2) Define $\rho_{\rm HS} = \frac{A^\dagger A}{{\rm tr}\{A^\dagger A\}}$.
(3) Obtain $U_{\rm HS}$ from the spectral decomposition $\rho_{\rm HS} = U_{\rm HS} D_{\rm HS} U_{\rm HS}^\dagger$ of diagonal $D_{\rm HS}$.

Thereafter, construct the new basis $\mathscr{B}_{\rm HS} = U_{\rm HS}\mathscr{B}_{\rm ref}$.

[1] I. Chuang and M. Nielsen, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] *Quantum State Estimation*, edited by M. G. A. Paris and J. Řeháček, Lecture Notes in Physics, Vol. 649 (Springer, Berlin, 2004).
[3] Y. S. Teo, *Introduction to Quantum-State Estimation* (World Scientific, Singapore, 2015).
[4] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, Diluted maximum-likelihood algorithm for quantum tomography, Phys. Rev. A **75**, 042108 (2007).
[5] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Quantum-State Reconstruction by Maximizing Likelihood and Entropy, Phys. Rev. Lett. **107**, 020404 (2011).
[6] J. Shang, Z. Zhang, and H. K. Ng, Superfast maximum-likelihood reconstruction for quantum tomography, Phys. Rev. A **95**, 062336 (2017).
[7] P. Busch and P. J. Lahti, The determination of the past and the future of a physical system in quantum mechanics, Found. Phys. **19**, 633 (1989).
[8] E. Prugovečki, Information-theoretical aspects of quantum measurement, Int. J. Theor. Phys. **16**, 321 (1977).
[9] A. Kalev, R. L. Kosut, and I. H. Deutsch, Quantum tomography protocols with positivity are compressed sensing protocols, npj Quantum Inf. **1**, 15018 (2015).
[10] C. H. Baldwin, I. H. Deutsch, and A. Kalev, Strictly-complete measurements for bounded-rank quantum-state tomography, Phys. Rev. A **93**, 052105 (2016).
[11] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Scalable multiparticle entanglement of trapped ions, Nature (London) **438**, 643 (2005).
[12] J. G. Titchener, M. Gräfe, R. Heilmann, A. S. Solntsev, A. Szameit, and A. A. Sukhorukov, Scalable on-chip quantum state tomography, npj Quantum Inf. **4**, 19 (2018).
[13] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory **52**, 1289 (2006).
[14] E. J. Candès and T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans. Inf. Theory **52**, 5406 (2006).
[15] E. J. Candès and B. Recht, Exact matrix completion via convex optimization, Found. Comput. Math. **9**, 717 (2009).
[16] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography Via Compressed Sensing, Phys. Rev. Lett. **105**, 150401 (2010).
[17] A. Steffens, C. A. Riofrío, W. McCutcheon, I. Roth, B. A. Bell, A. McMillan, M. S. Tame, J. G. Rarity, and J. Eisert, Experi-

mentally exploring compressed sensing quantum tomography, Quantum Sci. Technol. **2**, 025005 (2017).

[18] C. A. Riofrío, D. Gross, S. T. Flammia, T. Monz, D. Nigg, R. Blatt, and J. Eisert, Experimental quantum compressed sensing for a seven-qubit system, Nat. Commun. **8**, 15305 (2017).

[19] D. Ahn, Y. S. Teo, H. Jeong, F. Bouchard, F. Hufnagel, E. Karimi, D. Koutný, J. Řeháček, Z. Hradil, G. Leuchs, and L. L. Sánchez-Soto, Adaptive Compressive Tomography with no *a Priori* Information, Phys. Rev. Lett. **122**, 100404 (2019).

[20] M. Rani, S. B. Dhok, and R. B. Deshmukh, A systematic review of compressive sensing: Concepts, implementations and applications, IEEE Access **6**, 4875 (2019).

[21] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, Quantum tomography via compressed sensing: Error bounds, sample complexity and efficient estimators, New J. Phys. **14**, 095022 (2012).

[22] E. J. Candès and T. Tao, Decoding by linear programming, IEEE Trans. Inf. Theory **51**, 4203 (2005).

[23] E. J. Candès, J. K. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commun. Pure Appl. Math. **LIX**, 1207 (2006).

[24] B. Recht, M. Fazel, and P. Parrilo, Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization, SIAM Rev. **52**, 471 (2010).

[25] S. T. Flammia, A. Silberfarb, and C. M. Caves, Minimal informationally complete measurements for pure states, Found. Phys. **35**, 1985 (2005).

[26] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gómez, G. B. Xavier, G. Lima, and A. Delgado, Five Measurement Bases Determine Pure Quantum States on Any Dimension, Phys. Rev. Lett. **115**, 090401 (2015).

[27] F. Mezzadri, How to generate random matrices from the classical compact groups, Not. AMS **54**, 592 (2007).

[28] P. Ćwikliński, M. Horodecki, M. Mozrzymas, Ł. Pankowski, and M. Studziński, Local random quantum circuits are approximate polynomial-designs: Numerical results, J. Phys. A: Math. Theor. **46**, 305301 (2013).

[29] N. J. Russell, L. Chakhmakhchyan, J. L. O'Brien, and A. Laing, Direct dialing of Haar random unitary matrices, New J. Phys. **19**, 033007 (2017).

[30] L. Banchi, W. S. Kolthammer, and M. S. Kim, Multiphoton Tomography with Linear Optics and Photon Counting, Phys. Rev. Lett. **121**, 250402 (2018).

[31] K. Życzkowski and H.-J. Sommers, Hilbert–Schmidt volume of the set of mixed quantum states, J. Phys. A: Math. Gen. **36**, 10115 (2003).

[32] M. Kech and M. M. Wolf, Constrained quantum tomography of semi-algebraic sets with applications to low-rank matrix recovery, Inf. Inference **6**, 171 (2016).

[33] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).

[34] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, Certifying the restricted isometry property is hard, IEEE Trans. Inf. Theory **59**, 3448 (2013).

[35] D. N. Tran, S. Huang, S. P. Chin, and T. D. Tran, Low-rank matrices recovery via entropy function, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016* (IEEE, New York, 2016).

[36] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx, September 2013.

[37] M. C. Grant and S. P. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, Vol. 371, edited by V. D. Blondel, S. Boyd, and H. Kimura (Springer-Verlag, Berlin, Heidelberg, 2008), pp. 95–110.