Title:

Quantum certification and benchmarking

Author(s):

Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud & Elham Kashefi

Citation:

"Eisert, J., Hangleiter, D., Walk, N. et al. Quantum certification and benchmarking. Nat Rev Phys 2, 382–390 (2020). https://doi.org/10.1038/s42254-020-0186-4"

Quantum certification and benchmarking

J. Eisert,[1,2,3,*] D. Hangleiter,[1,*] N. Walk,[1] I. Roth,[1] D. Markham,[4,5] R. Parekh,[4,5] U. Chabaud,[4,5] and E. Kashefi[4,5,6]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Department of Mathematics and Computer Science, Freie Universität Berlin, 14195 Berlin, Germany*
[3]*Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany*
[4]*Paris Center for Quantum Computing, CNRS FR3640, Paris, France*
[5]*Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, F-75005 Paris, France*
[6]*School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, United Kingdom*

Concomitant with the rapid development of quantum technologies, challenging demands arise concerning the certification and characterization of devices. The promises of the field can only be achieved if stringent levels of precision of components can be reached and their functioning guaranteed. This review provides a brief overview of the known characterization methods of certification, benchmarking, and tomographic recovery of quantum states and processes, as well as their applications in quantum computing, simulation, and communication.

a. Introduction. Recent years have seen a rapid development of quantum technologies, promising new real-world applications in communication, simulation, sensing and computation [4]. Quantum internet infrastructure enables unconditionally secure transmission and manipulation of information [70, 124]. Highly engineered quantum devices allow for the simulation of complex quantum matter [29]. While noisy intermediate scale quantum devices [94] are on the verge of outperforming classical computing capabilities [7], a longer term perspective of fault tolerant quantum computers [23] aims to solve impactful problems from industry that are out of reach for classical computers. These prospects come along with enormously challenging prescriptions concerning the precision with which the components of the quantum devices function. The task of ensuring the correct functioning of a quantum device in terms of the accuracy of the output is referred to as *certification* or sometimes verification. *Benchmarking* more generally assigns a reproducible performance measure to a quantum device.

The very tasks of certification and benchmarking are challenged by intrinsic quantum features: The involved configurations spaces have enormous dimensions, a serious burden for any characterization. What is more, certification comes along with an ironic twist: It is highly non-trivial in light of the fact that certain quantum computations are expected to exponentially outperform any attempt at classically solving the same problem. While a large-scale universal quantum computers are still out of reach, already today do we have access to quantum simulators, that is, special-purpose, highly controlled quantum devices aimed at simulating physical systems [13, 29]. And, indeed, for such systems, often, no efficient classical simulation algorithm is available. As a consequence, as quantum devices are scaled up to large system sizes, application-specific tools of certification are required that go beyond standard approaches such as re-simulating a device on a classical computer or full tomographic reconstruction. It is such specifically 'quantum' certification tools that this review summarizes and puts into context.

To do so, we offer a framework in which the resource cost, the information gained as well as the assumptions made in such approaches are cast very naturally. We then turn to charting the landscape of different approaches to quantum certification within our framework, ranging from practically indispensable, economic diagnostic tools such as randomized benchmarking to cryptographically secure techniques such as self-testing or the verification of arbitrary quantum computations in an interactive fashion [43, 50]. In doing so, we aim at painting a panoramic sketch of this landscape useful for categorizing various tools and putting them into context. Some of the methods we lay out are crucial for the development and engineering of noisy near-term devices, some will find practical applications once large-scale sophisticated devices become available. The main importance of yet others rests in setting the stage of the possible, highlighting extremal points of this landscape, and inviting future method development to find good compromises between desirable features of a protocol. In this review, we therefore aim to be explicit about the resource costs and assumptions made in specific protocols.

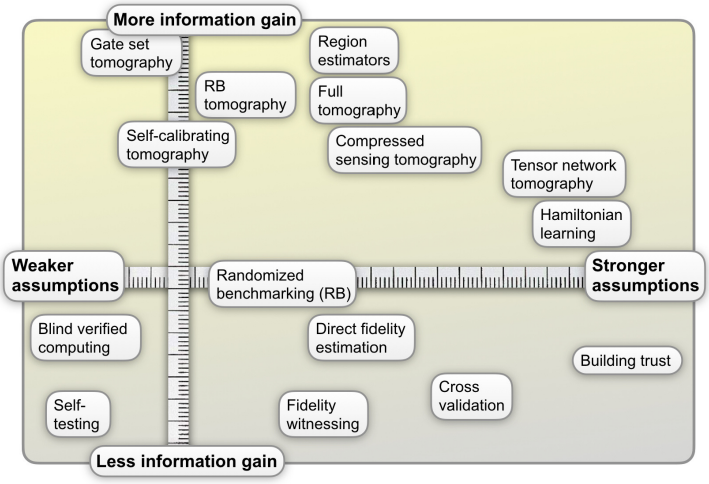

FIG. 1. Schematic of a classification scheme for some of the certification protocols discussed in this article. One axis quantifies the information gain, the other the strength and number of assumptions required. For clarity, we leave out the complexity of the protocol. An example of a protocol is discussed in BOX 2.

* J. E. and D. H. have contributed equally.

BOX 1: Measures of quality

Since a certificate should guarantee the correct functioning of a given quantum process or the correct preparation of a desired quantum state, it should ideally be phrased in terms of a measure of distance between two such objects that has an operational interpretation as their worst-case distinguishability. Preferably, such measures should also be *composeable*, meaning that individual device or protocol certificates can be combined to certify larger, composite systems, which is especially crucial for cryptographic applications [93]. Specifically, certificates for quantum states are often phrased in terms of the *trace distance* $d(\sigma,\rho) = \mathrm{tr}[|\sigma-\rho|]/2$, and for quantum channels in terms of the *diamond norm* [123], which can be conceived as a stabilized trace distance for channels. For state vectors $|\psi\rangle$, certificates can be easily phrased in terms of the fidelity $F(\rho,|\psi\rangle\langle\psi|) = \langle\psi|\,\rho\,|\psi\rangle$, which measures the overlap between $|\psi\rangle\langle\psi|$ and ρ. The *quality of quantum gates* is commonly expressed in terms of the *average gate fidelity* [105]. This quantity expresses the overlap of the output with the anticipated output of the quantum gate, in a way that is agnostic to the direction in Hilbert space. The fidelity for quantum states directly bounds the trace distance via $1-F(\sigma,\rho)^{1/2} \leq d(\sigma,\rho) \leq (1-F(\sigma,\rho))^{1/2}$ and therefore certifies *worst-case* performance. In contrast, the average gate fidelity only yields useful bounds that do not incur a dimension factor for the diamond norm in certain special situations [75] and can therefore certify performance typically only *on average*. For concrete tasks at hand, other measures of quality may apply. Building upon such notions of fidelities, specific measures of quality have been introduced in different contexts. Examples include the cross-entropy [18] or cross-entropy benchmarking [7] aimed at verifying classical distributions or the *quantum volume* [32] aimed at capturing the quality of entire quantum circuits or gate sets. In the context of quantum simulation, as breaking up the entire scheme into physical and conceptual building blocks is less obvious, notions of coherence [109], entanglement [55], non-classicality [81] or purity (i.e., $\mathrm{tr}(\rho^2)$) are made use of. In order to compare the quality of devices, one can define a precisely reproducible task and take a well-defined measure of performance in this task (e.g., number of secure bits in a quantum key distribution protocol [103]) as a valid figure of merit, a benchmark, in itself.

b. Classifying quantum certification. In any task of quantum certification, the core aim is to establish the correct functioning of a quantum device. Given the enormous effort of a full tomographic characterization of quantum states and processes, in many practical applications, protocols for certification will necessarily be constrained in the available resources and at the same time governed by the advice one hopes to gain from the protocol. With this in mind, it is instructive to conceptualize the quantum certification problem as a protocol between the *quantum device*, seen as being powerful, and its *user*, who is restricted in her or his measurement devices and computational power. One can classify schemes according to the effort and information gained, as well as in the assumptions made on the device and its user (see Fig. 1). Ultimately, even when one aims for tomographic knowledge, one may conceive of certification as a protocol that outputs 'accept' if the device functions correctly, and 'reject' if it does not. Whether the protocol accepts or rejects is determined according to reasonable *measures of quality* that are appropriate for the respective property of the device being certified (BOX 1).

The *assumptions* made on the devices and their users depend, among other aspects, on one's trust levels and the specific setting at hand. Typically, which assumptions are made also has an effect on the protocol's complexity in one way or the other, or even renders certification feasible in the first place. Conceptually speaking, there are three building blocks entering, each of which equipped with certain assumptions. This is, firstly, the *quantum device* to be certified, a distinct and often physically separate entity. In experimental scenarios it is commonly reasonable to include knowledge concerning the underlying physical mechanism and potential sources of error in terms of an adequate modelling framework. However, it can also make sense to merely assume that the device is a quantum mechanical object. Secondly, it is useful to distinguish the *quantum measurement apparata* used in the characterization, which might include state preparation and short circuits. In an idealized setting, they may be assumed to be perfect. More practically relevant are situations in which one has a solid understanding of their functioning and characterized their efficiencies to some level. In many physical architectures, in particular in key platforms for quantum simulation, one can perform certain quantum measurements very accurately, but is severely limited in the *type of measurements* that can be performed. Thirdly, and finally, there is the *classical data processing* which consumes storage capacity and processing time. Ultimately, any characterization provides classical numbers. The *device-independent setting* makes no assumptions at all about the measurement apparata and the device, taking into account the final data only. Such a setting is adequate, for example, if the device is a remote and untrusted quantum device that may be accessible only through the cloud.

The effort or *complexity* of such a certification protocol can be divided into several distinct parts: This is the number of different settings or rounds in which data is obtained from the measurement device *(measurement complexity)*. Implementing those different settings might require different *quantum computational effort* as for example quantified by the length of the circuit that implements a certain measurement. Then, there is a minimal number of experiments and resulting samples that need to be obtained for a protocol to meaningfully succeed *(sample complexity)*. Finally, one needs to process those samples involving classical computational effort in time and space *(post-processing complexity)*.

Often, the complexity of a protocol can be traded for the *amount of information* about the device that the user can extract when running the protocol. Such information is crucial when it comes to designing and improving a concrete experimental setup, while it may be less important when the user's goal is merely to check the correct functioning of, say, a newly bought device, or a remote server. Indeed, whilst many cer-

tification techniques have been developed with specific applications in mind, the abstract criteria outlined above provide a framework to discuss the strengths, weaknesses and relevant application of these techniques in more general terms.

For example, one can consider the relative importance of these criteria for applications on the spectrum from exploratory science, through proof-of-principle demonstrations to large scale technological implementation. At the exploratory end of the spectrum, information gain is at a premium as the researcher endeavors to maximise their understanding of the underlying physics. Often, such experiments are small-scale and involve well characterised measurement devices probing a relatively less understood target device. Here, the complexity of a technique will be less important and, while some assumptions regarding the measurement apparatus may be reasonable, they should be avoided as much as possible regarding the device to be certified. As we will see in the next section, this combination of desiderata would motivate the use of quantum state tomography and related techniques. In a proof-of-principle demonstration of a larger-scale but better controlled device the relative importance of the assumptions made and the information gained is reduced with respect to complexity involved. If presented instead with a high-quality, large-scale device, efficiency will become crucial and remote users may prioritise simple-to-use certification techniques such as self-testing or benchmarking at the level of applications rather than hardware.

We now present and assess various tools for characterisation ordered according to, first, the information that may be extracted from the protocol, and, second, the assumptions made in the protocol. Rather than being exhaustive and technically detailed, our selection highlights distinct points within our framework with the goal to sketch a panoramic view of the landscape it gives rise to. In addition to the main text, we provide a tabular overview in which we quantitatively assess exemplary certification protocols for applications in cloud computing, demonstrating a quantum advantage, and quantum simulation and computation according to our classification (TABLE I). We illustrate how to read this table by means of exemplary cases in BOX 4.

c. Certification protocols. In many scenarios, it is reasonable to assume that one's quantum measurements are rather well characterized and that the object of interest is either a quantum state or process that can be accessed in independently identically distributed (i.i.d.) experiments. These assumptions are often very natural in laboratory settings in which the quantum device can be directly accessed. They are therefore at the heart of many characterization protocols and shall be our starting point for now.

The most powerful but at the same time most resource-intense such technique of certification is full *quantum tomography* [64, 67]. Here, the idea is to obtain knowledge of the full quantum state or process by performing sufficiently many (trusted) measurements. Given tomographic data, one can in particular obtain a certificate that the state lies in some region in state space. For many years such regions were typically constructed heuristically by first applying maximum likelihood estimation to construct a point estimate of the state [65] and then using resampling techniques to obtain error bars. More recently, techniques to obtain more rigorous region estimates have appeared including *Bayesian credibility regions* [14, 41] and *confidence regions* [15, 28, 122] where the former are usually smaller but depend strongly upon the Bayesian prior. Most importantly, from these tomographic reconstructions, one exactly learns the nature of deviation of the imperfect implementation to the target. Such data proves crucial when designing experimental setups as it yields information about the particular sources of errors present in the setup and hence functions as 'actionable advice' on how to improve the setup.

However, generic quantum state and process tomography is excessively costly in the size of the quantum system. Fortunately, many quantum states and processes that are encountered in realistic experiments exhibit significant structure: States are often close to being pure or have approximately low rank, so that methods of *compressed sensing tomography* [53, 56, 69] can be applied in which less resource expensive or more reliable recovery is possible based on the same type of (but randomly chosen) measurements compared to full tomography. Similarly, quantum processes are often close to being unitary [45, 72]. For local Hamiltonian systems, even further structure of locality comes into play. In particular, tensor network states can provide meaningful variational sets for *tensor network tomography*, which basically

BOX 2: Randomized benchmarking

Randomized benchmarking (RB) refers to a collection of methods that aim at reliably estimating the magnitude of an average error of a quantum gate set in robust fashion against *state preparation and measurement (SPAM)* error. It achieves this goal by applying sequences of feasible quantum gates of varying length, so that small errors are amplified with the sequence length leading. From a pragmatic point of view, RB protocols thereby define benchmarks that can be used to compare different digital quantum devices. In important instances, the benchmark can be related to the average gate fidelity, rendering RB protocols flexible certification tools. To this end, a group structure of the gate set is made use to achieve two goals: On the one hand, this is to control the theoretical prediction of error-free sequences. On the other hand, this allows one to analyze the error contribution after averaging using representation theory. Originally devised for random unitary gates [34, 38, 76], RB is most prominently considered for Clifford gates [73, 78], and has been extended to other finite groups [9, 24, 33, 39, 62, 90]. Assumptions on having identical noise levels per gate have been lessened [118], and RB with confidence introduced [61, 121]. RB schemes have been generalized to other measures of quality, such as relative average gate fidelities [79] with specific target gates, fidelities per symmetry sector [24, 90], the unitarity [117], measures for losses, leakage, addressibility and cross-talk [49, 119, 120] or even tomographic schemes that combine data from multiple RB experiments [47, 71, 102]. In addition, RB protocols have been devised that directly work on the level of generating gate sets [48, 95].

BOX 3: Certifying a quantum advantage or quantum computational supremacy

Using a quantum computer to efficiently perform computational tasks that are provably intractable for classical computers marks a key milestone in the development of quantum technologies. Various sub-universal models of quantum computing have been proposed to demonstrate, with near-term achievable technology, a so-called *quantum advantage* or *quantum computational supremacy*. A crucial part of the demonstration of this claim with a given model is the verification of the output of the corresponding quantum device. But the nature of the computational task is precisely such that it cannot be reproduced classically and therefore the traditional means of verifying a computation fail. What is more, the proposed sub-universal quantum devices produce samples from exponentially flat probability distributions to the effect that it requires exponentially many samples to classically verify that the obtained samples are indeed distributed according to the target distribution, independently of the hardness of producing the samples [58, 116]. The latter result severely restricts the possibilities for deriving classical verification protocols for quantum computational supremacy even under the assumption that the verifier has access to *arbitrary computational power*.

To circumvent this no-go result and arrive at a sample-efficient verification protocol one may take very different routes: First, one might ask for less than verification of the full output distribution such as merely distinguishing against the uniform or certain efficiently sampleable distributions, which can often be done in a computationally efficient way [2, 25, 92, 108]. Allowing for exponential time in classical post-processing, one can also sample-efficiently verify coarse-grained versions of the target distribution [19], make use of certain complexity-theoretic assumptions [3], or assumptions on the noise in the quantum device [18, 19, 40]. The latter allows one to use weaker measures, like the *cross-entropy* [18, 19] or variants thereof such as the cross-entropy benchmarking fidelity [7]. If one gives qualitatively more power to the user, e.g., trusted single-qubit measurements [59, 91, 112], this even allows one to fully efficiently verify the prepared quantum state and thereby the sampled distribution. Finally, one may use more complicated, interactive protocols which require a universal quantum device, e.g., the one presented in Ref. [80], which relies on the post-quantum security of a certain computational task to classically delegate a universal computation. Given the importance of verifying a quantum advantage, it is a pressing challenge to derive fully efficient verification protocols which involve minimal assumptions. We expect that this will require custom-tailored techniques for the different available proposals.

makes the structural assumption that there is little entanglement in the state, an assumption that is often valid for quantum many-body states to an extraordinarily good approximation [10, 31, 66, 89]. Also, variational sets inspired by *machine learning* have been considered [26, 113]. In such situations, the effort of quantum state and process tomography can be significantly reduced. At least for intermediate-sized systems, such techniques are practically highly important.

If one is only interested in certain properties of a quantum state or process one may resort to so-called *learning techniques*, which scale much more favourably. For instance, one may merely be interested in *probably approximately correctly (PAC)* learning the expected outcomes of a certain set of measurements, e.g., local observables on the quantum state. PAC learning is possible with a measurement complexity that scales only linearly (in the number of qubits) in certain settings [1, 100, 101] but still incurs exponential computational effort. In another instance of learning, one might be confident that the given data is described by a certain restricted Hamiltonian (or Liouvillian) model whose parameters are however not known. *Hamiltonian (or Liouvillian) learning techniques* solve this task and recover the Hamiltonian parameters from suitable data [52, 63].

In contrast to the aforementioned tools for characterizing a quantum device, *fidelity estimation* aims merely at determining the overlap of the actual quantum state or process implemented in a given setup with the ideal one. While fidelity estimation yields much less information than full tomography, one saves tremendously in measurement and sample complexity. In fact, using *importance sampling* one can estimate the fidelity of an imperfect preparation of certain pure quantum states in constant measurement complexity [46]. This protocol can be extended to optimally estimating the fidelity of quantum channels [98].

A yet weaker notion than fidelity estimation is *fidelity witnessing*. The idea of a fidelity witness is to cut a hyper-plane through quantum state space which separates states close in fidelity to a target state from those far away. Efficient fidelity witnesses can often be derived in settings in which the target state satisfies some extremality property so that it lies in a low-dimensional corner of state space, such as certain multipartite entangled states [91], Gaussian bosonic states [6] or ground states of local Hamiltonians [59].

A still weaker approach merely aims at verifying or estimating the presence of certain key properties, such as entanglement from realistic measurements, to, say, observe entanglement propagation [68]. Here again, notions of (quantitative) witnesses that provide bounds to entanglement measures play an important role [8, 36, 54]. Such witnesses can be measured by exploiting randomness [22].

In case one has a good understanding of the physical mechanisms governing the device, it is often useful to *build trust* in the quantum device. This approach is particularly prominent in the context of quantum simulation: Here, the idea is to certify a quantum device by validating its correct functioning in certain classically simulable regimes through comparison to classical simulations [20, 104, 114, 115]. In some instances, stronger statements can be made when invoking notions of *self-validation* [74] or *cross-platform verification* [37]. It is also common to certify the components of a device, for example, individual gates, and extend the trust obtained in this way to the full device, making the assumption that all sources of errors are already present for the individual components. In such approaches, it is assumed that no additional sources of errors arise when moving out of the strictly certifiable regime again.

An important drawback of most schemes discussed so far, however, is that they assume i.i.d. state preparations. This lim-

BOX 4: A guide to TABLE 1

In TABLE 1, we are comparing a wide range of protocols some of which are structurally distinct. We have settled on certain criteria described above to meaningfully compare a variety of techniques in a unified language. But of course more fine-grained distinctions are necessary to exhaustively describe all the protocols. For example, in delegated computing protocols natural figures of merit include the number of communication rounds and transmitted bits as well as overhead in terms of qubit number. However, such quantities do not appear in state tomography protocols. As the number of qubits and rounds affects the number of single-qubit measurements that need to be performed by the server in a blind computing protocol our criteria capture the effort required to perform the protocol. Let us give provide two examples.

Example 1 (Blind computing via trap qubits [44]): Here, the client prepares and sequentially transmits to the server a product quantum state of $N \in O(nD\log(1/\epsilon)$ many qubits in order to delegate and verify a depth-D quantum computation on n qubits with trace-distance error ϵ. The server entangles the qubits in a graph state, measures all of them, and sends the outcomes to the client for post-processing who simply compares certain outcomes. The number of distinct single-qubit measurements is therefore given by one choice of settings for a single N-qubit measurement, while the number of samples is 1 as the protocol is single-shot. To obtain the certificate, $O(N)$ many of the single-qubit measurement outcomes need to be compared and hence the post-processing is linear in that number.

Example 2 (Low-rank state tomography with 2-designs [57]): Here, one repeatedly measures a positive operator valued measure (POVM) that constitutes a complex projective 2-design on $O(2^n r^2/\epsilon^2)$ i.i.d. copies of an n-qubit rank-r quantum state. Such a POVM consists of at least $O(4^n)$ elements and, hence, requires an exponential number of observables. The sampling complexity is of order of the degrees of freedom of a rank-r state up to an additional factor of r. From the frequency estimates one calculates a linear least-square estimator and subsequently projects the result onto quantum states which requires a time complexity of $O(2^{3n})$ on a classical computer. A trace-norm ball around the obtained estimate is a confidence region depending only on the estimate's rank.

itation can be overcome using *quantum de Finetti arguments* to obtain non-i.i.d. tomographic regions [28, 122] and distance certificates [112], the use of which has been optimized in various works for the case of *graph states* [82, 111] as well as for *continuous variable states* [27].

More severely still, in the standard setting a high level of trust in the measurement devices is required giving rise to a vicious cycle: After all, to calibrate the measurement devices in the first place, one requires quantum probe states which are well characterized, a task that requires well calibrated measurement devices. This raises the question whether one can simultaneously learn about the quantum device and the quantum measurement apparatus in a self-consistent or semi-device-dependent way. The rather extreme and resource-intense solution to this problem is *gate set tomography* which instead of focusing on a single quantum channel or state, characterizes an entire set of quantum gates, the state preparation and the measurement self-consistently from different gate sequences [16, 17, 85]. Other solutions have been demonstrated in optics settings where one can perform state tomography in a *self-calibrating way* [21, 86]. Such schemes at times even come with error bars [107]. One can also exploit well characterized reference states such as coherent states [87] as a lever to perform uncalibrated tomography [97]. In another vein, one can mitigate uncertainty in the model that generated the data by using model averaging techniques [42]. A particularly important example of fidelity-estimation protocols for quantum processes that break this vicious cycle have been proposed in the context of *randomized benchmarking* [34, 38, 76, 78] (see (BOX 2)).

If a quantum device already allows for a level of abstraction that is close to an envisioned application, one can use more specific or even ad hoc performance measures in simple tasks as a benchmark. For example, quantum key distribution is usually certified entirely at the application level through the number of secure distributed bits [103]. One can also run small instances of quantum algorithms on prototypes of digital quantum computers [35, 77].

One could imagine, however, applications where even mild assumptions cannot be guaranteed. In such a scenario, one could utilize a range of cryptographic tool-kits to ensure that the above assumptions are indeed enforced. One prominent example of such setting is where one has to work in a black-box setting, i.e., with no assumptions made about the underlying devices. Remarkably, the non-local correlations demonstrated by quantum mechanics allow for certain entangled states and non-commuting measurements to be certified in this setting (up to local isometries) solely via the observed statistics [83]. This procedure of *self-testing* is typically achieved through the violation of a Bell inequality, with the paradigmatic example being the maximal violation of the CHSH inequality which self-tests non-commuting Pauli measurements made upon a maximally entangled pair of qubits. A substantial body of literature has extended these results in many directions, including generalisations to multi-partite entangled states [84], gates and instruments [106], and approximate collections of maximally entangled qubit pairs [99] which have been made increasingly robust [88] (a recent and comprehensive review of self-testing can be found in Ref. [110]).

In the context of computation, the key idea of device-independent schemes is to *hide* the delegated computation from a remote black box server in such a way that the powerful server cannot retrieve any information about the computation without leaking to the client that a deviation has occurred. The use of the *quantum twirling lemma* or similar technique allows one to simplify the analysis under a general deviation (with no assumptions) to a simple (i.i.d.) case leading to efficient *verified blind quantum computation* schemes [30, 44, 99]. Guarantees of correctness have been achieved in this manner in various scenarios, giving different degrees

of control to the user. These powerful verification schemes, while removing many trust assumptions and providing efficient protocols, remain only applicable to a remote verifier with limited quantum capacity, such as single qubit gates [44], or access to entangled servers [30, 99]. These last obstacles have recently been overcome by utilizing yet another cryptographic toolkit, this time from the classical domain. The usage of post-quantum secure collision-resistant hash functions has enabled a fully classical client to hide and verify the remote quantum computation [51, 80]. However, these new schemes come with a significant overhead that can be reduced to some extent [5], and they are no longer fully unconditionally secure, as they are based on a computational assumption, that is, the existence of classical problems that are computationally hard to solve even for a quantum computer [96]. We provide a case study of how different certification methods can be applied in the context of verifying a quantum computational advantage (BOX 3).

d. Outlook. In this review, we have provided an overview of methods for certifying and benchmarking quantum devices as they are increasingly becoming of key importance in the emerging quantum technologies (for detailed, up-to-date information, see (BOX 5)). We hope that our framework will prove to be a useful means of assessing and putting into a holistic context methods to be developed in the future. Indeed, achieving good compromises between resource cost, obtainable information and assumptions made in a protocol may well be a make-or-break topic for quantum technologies. In this mindset, this review – and the quantitative framework provided here – is also meant to be an invitation and guideline for future method development to a growing field of research that combines sophisticated mathematical reasoning with a data-driven experimental mindset.

e. Acknowledgements. We gratefully acknowledge discussions with D. Gross and J. Helsen, in addition to many other members of the scientific community. J. E. acknowledges funding from the DFG (CRC 183 project B01, EI 519/9-1, EI 519/14-1, EI 519/15-1, MATH+, EF1-7), the BMWF (Q.Link.X), the BMWi (PlanQK), and the Templeton Foundation. N. W. acknowledges funding support from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 750905. This work has also received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No. 817482 (PASQuanS). E. K. and D. M. acknowledge funding from the ANR project ANR-13-BS04-0014 COMB, E. K. by the EPSRC (EP/N003829/1).

BOX 5: Online certification library

Some of the present authors have curated an online library of certification protocols on the Quantum Protocol Zoo, hosted at wiki.veriqloud.fr under certification library. The aim of this repository is to provide a compact and precise review of the existing certification techniques and the corresponding protocols beyond the scope of this work. As of today, this library consists of a few concrete protocols in a specified format, classified in different techniques, where the technique page also includes a brief description, properties and the references. For every protocol, we provide its detailed outline, the assumptions considered, resources and requirements, a mathematical description of the procedure and properties including sample and measurement complexity. Certification plays an important role in the development of quantum devices and we hope this library will help the community classify the certification techniques and also keep updating them with the progress in this field.

[1] S. Aaronson. The learnability of quantum states. *Proc. Royal Soc. A*, 463:3089–3114, 2007.

[2] S. Aaronson and A. Arkhipov. Bosonsampling is far from uniform. 2013. arXiv:1309.7460.

[3] S. Aaronson and L. Chen. Complexity-theoretic foundations of quantum supremacy experiments. 2016. arXiv:1612.05903.

[4] A. Acin, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm. The European quantum technologies roadmap. *New J. Phys.*, 20:080201, 2018.

[5] G. Alagic, A. M. Childs, and S.-H. Hung. Two-message verification of quantum computation. *arXiv:1911.08101 [quant-ph]*. arXiv: 1911.08101.

[6] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert. Reliable quantum certification for photonic quantum technologies. *Nature Comm.*, 6:8498, 2015.

[7] F. Arute and et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574:505–510, 2019.

[8] K. M. R. Audenaert and M. B. Plenio. When are correlations quantum? *New J. Phys.*, 8:266, 2006.

[9] R. Barends, J. Kelly, A. Veitia, A. Megrant, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, E. Jeffrey, C. Neill, P. J. J. O'Malley, J. Mutus, C. Quintana, P. Roushan, D. Sank, J. Wenner, T. C. White, A. N. Korotkov, A. N. Cleland, and John M. Martinis. Rolling quantum dice with a superconducting qubit. *Phys. Rev. A*, 90:030303, 2014.

[10] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio. Scalable reconstruction of density matrices. *Phys. Rev. Lett.*, 111:020401, 2013.

[11] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio. A scalable maximum likelihood method for quantum state tomography. *New J. Phys.*, 15:125004, 2013.

[12] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert. Architectures for quantum simulation showing a quantum speedup. *Phys. Rev. X*, 8:021010, 2018.

[13] I. Bloch, J. Dalibard, and S. Nascimbene. Quantum simulations with ultracold quantum gases. *Nature Phys.*, 8:267, 2012.

[14] R. Blume-Kohout. Optimal, reliable estimation of quantum states. *New J. Phys.*, 12:043034, 2010.

Application	*Verification protocol*	*Information*	*Complexity*				*Assumptions*
			Measurements	Feasibility	Samples	Post-processing	
Quantum networks and cloud computing							
Cloud quantum computing	Blind computing via trap qubits [44]	Trace distance	$1\times$ $O(nD\log(1/\epsilon))$	single-qubit preparation	1	$O(nD\log(1/\epsilon))$	Client prepares and sends single qubits, receives classical data
	Blind computing via Bell test [30]	Trace distance	$O(D)\times$ $\tilde{O}(nD\log(1/\epsilon))$	classical	1	$O(nD\log(nD)$ $\cdot\log(1/\epsilon))$	Non-communicating and spatially separated servers
	Classically verified quantum computing [80]	Trace distance	$2\times$ $O(n^3D^2\log(1/\epsilon)$	classical	1	$O(n^3D^2\cdot$ $\cdot\log(1/\epsilon))$	*Learning With Errors* problem is hard for quantum computers
Secret sharing	Graph state schemes [82]	Trace distance	$O(1/\epsilon)\times n$	Pauli obs.	1	$O(n/\epsilon)$	i.i.d., trusted single-qubit measurements for honest players
Quantum advantages							
Boson sampling	State discrimination [2, 25, 108]	Discriminate from uniform	$1\times n$	classical	$O(1)$	$O(n)$	Characterized random unitary
Random circuit sampling	HOG verification [3]	Classically hard task achieved	$1\times n$	classical	$O(1)$	$O(\exp(n))$	Generating larger-than-median outputs is classically hard.
	Cross-entropy verification [18, 19]	TV distance	$1\times n$	classical	$O(n^2/\epsilon^2)$	$O(\exp(n))$	$H(p_{\text{device}})\geq H(p_{\text{ideal}})$
	Identity testing [58, 116]	TV distance	$1\times n$	classical	$O(\sqrt{2^n}/\epsilon^2)$	$O(\exp(n))$	None
Shallow circuits	Local measurements [12, 59]	Fidelity witness	$2\times n$	Hadamard and T gate	$O(n^2/\epsilon^2)$	$O(n^3/\epsilon^2)$	Trusted single-qubit measurements
Quantum computation and simulation							
Fault-tolerant QC	Randomized benchmarking [60]	Av. gate fidelity	$1\times n$	standard basis	$O(1/\epsilon_R^2)$	$O(1/\epsilon_R^2)$	i.i.d., SPAM robust
	Gate set tomography [16]	Point estimates	$1\times n$	standard basis	?	?	i.i.d., fully self-consistent
State preparation	Maximum-likelihood tomography [64]	Point estimate	$\exp(O(n))\times n$	?	?	?	i.i.d., trusted measurements
	Error-bar tomography [122]	Confidence polytope	$\exp(O(n))\times n$	?	?	?	i.i.d., trusted measurements
	Rank-r tomography with 2-designs [56]	Point estimate + confidence cert.	$\exp(O(n))\times n$	?	$\tilde{O}(2^nr^2/\epsilon^2)$	$O(2^{3n})$	i.i.d., trusted measurements
	Direct fidelity estimation [46]	Fidelity estimate	$O(1/\epsilon^2)\times n$	Pauli obs.	$O(2^n/\epsilon^2)$	$O(2^n/\epsilon^2)$	i.i.d., trusted measurements
	Cross-platform verification [37]	Fidelity est. on n_a-qubit subsys.	$O(e^{bn_a})\times n_a$, $b\lesssim 1$	arb. 1-qubit unitaries	$O(e^{bn_a})$	$O(e^{bn_a})$	i.i.d., trusted measurements
	Self-testing n Bell pairs [88]	Fidelity witness	$2\times n$	2 conjugate bases	1	$O(n)$	Non-communicating and spatially separated observers
Dynamical quantum simulation	MPO tomography [11, 63]	Point estimate + fidelity witness	$O(n)\times n$	local operator basis	$O(n^3)$	$\text{poly}(n)$	i.i.d., trusted measurements, MPO description
	Building trust [20, 104, 114, 115]	Trust	?	?	?	?	Plenty

TABLE I. We assess exemplary certification, characterization and benchmarking protocols for selected applications according to our framework: the information or certificate that can be obtained from the protocol, the effort or complexity divided into its different parts, and the assumptions made. We do so with respect to the number of qubits or optical modes, n, the depth of a circuit, D, and additive ϵ or relative ϵ_R error tolerances. We write the measurement complexity as the number of distinct measurements $\times$ the number of single qubits or optical modes those measurements act on. The sample complexity is the total number of samples required for the respective benchmark or certificate, we denote the Shannon entropy by H and abbreviate 'total-variation distance' by 'TV distance'. We refer to BOX 4 for guidance to the table. Where to the best of our knowledge no explicit resource scaling is in the literature we have put ?-marks.

[15] R. Blume-Kohout. Robust error bars for quantum tomography. 2012. arXiv:1202.5270.

[16] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit. 2013. arXiv:1310.4492.

[17] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature Comm.*, 8:14485, 2017.

[18] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven. Characterizing quantum supremacy in near-term devices. *Nature Phys.*, 14:595, 2018.

[19] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Phys.*, 15:159–163, 2019.

[20] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider. Emergence of coherence and the dynamics of quantum phase transitions. *PNAS*, 112:3641–3646, 2015.

[21] A. M. Braczyk, D. H. Mahler, L. A. Rozema, A. Darabi, A. M. Steinberg, and D. F. V. James. Self-calibrating quantum state tomography. *New J. Phys.*, 14:085003, 2012.

[22] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos. *Science*, 364:260, 2019.

[23] E. T. Campbell, B. M. Terhal, and C. Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549:172–179, 2017.

[24] A. Carignan-Dugas, J. J. Wallman, and J. Emerson. Characterizing universal gate sets via dihedral benchmarking. *Phys. Rev. A*, 92:060302, 2015.

[25] J. Carolan, J. D. A. Meinecke, P. Shadbolt, N. J. Russell, N. Ismail, K. Wörhoff, T. Rudolph, M. G. Thompson, J. L. O'Brien, J. C. F. Matthews, and A. Laing. On the experimental verification of quantum complexity in linear optics. *Nature Phot.*, 8:621–626, 2014.

[26] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita. Reconstructing quantum states with generative models. *Nature Mach. Intel.*, 1:155–161, 2019.

[27] U. Chabaud, T. Douce, F. Grosshans, E. Kashefi, and D. Markham. Building trust for continuous variable quantum states. 2019. arXiv:1905.12700.

[28] M. Christandl and R. Renner. Reliable quantum state tomography. *Phys. Rev. Lett.*, 109:120403, 2012.

[29] J. I. Cirac and P. Zoller. Goals and opportunities in quantum simulation. *Nature Phys.*, 8:264, 2012.

[30] A. Coladangelo, A. Grilo, S. Jeffery, and T. Vidick. Verifier-on-a-leash: New schemes for verifiable delegated quantum computation, with quasilinear resources. 2017. arXiv: 1708.07359.

[31] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu. Efficient quantum state tomography. *Nat. Comm.*, 1:149, 2010.

[32] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta. Validating quantum computers using randomized model circuits. 2018. arXiv:1811.12926.

[33] A. W. Cross, E. Magesan, L. S. Bishop, J. A. Smolin, and J. M. Gambetta. Scalable randomized benchmarking of non-clifford gates. *npj Quant. Inf.*, 2:16012, 2016.

[34] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, 2009.

[35] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe. Demonstration of a small programmable quantum computer with atomic qubits. *Nature*, 536(7614):63–66, 2016.

[36] J. Eisert, F. G. S. L. Brandao, and K. M. R. Audenaert. Quantitative entanglement witnesses. *New J. Phys.*, 9:46, 2007.

[37] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. Joshi, R. Blatt, C. F. Roos, and P. Zoller. Cross-platform verification of intermediate scale quantum devices. 2019. arXiv: 1909.01282.

[38] J. Emerson, R. Alicki, and K. Zyczkowski. Scalable noise estimation with random unitary operators. *J. Opt. B*, 7:S347–S352, 2005.

[39] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt. Characterizing large-scale quantum computers via cycle benchmarking. Nature Comm.(5347), 2019.

[40] S. Ferracin, T. Kapourniotis, and A. Datta. Verifying quantum computations on noisy intermediate-scale quantum devices. 2018. arXiv:1811.09709.

[41] C. Ferrie. High posterior density ellipsoids of quantum states. *New J. Phys.*, 16:023006, 2014.

[42] C. Ferrie. Quantum model averaging. *New J. Phys.*, 16:093035, 2014.

[43] J. F. Fitzsimons. Private quantum computation: An introduction to blind quantum computing and related protocols. *npj Quant. Inf.*, 3:23, 2017.

[44] J. F. Fitzsimons and E. Kashefi. Unconditionally verifiable blind computation. *Phys. Rev. A*, 96, 2017.

[45] 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.

[46] S. T. Flammia and Y.-K. Liu. Direct fidelity estimation from few Pauli measurements. *Phys. Rev. Lett.*, 106:230501, 2011.

[47] S. T. Flammia and J. J. Wallman. Efficient estimation of Pauli channels. 2019. arXiv:1907.12976.

[48] D. S. Franca and A.-L. Hashagen. Approximate randomized benchmarking for finite groups. *J. Phys. A*, page 395302, 2018.

[49] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.*, 109:240504, 2012.

[50] A. Gheorghiu, T. Kapourniotis, and E. Kashefi. Verification of quantum computation: An overview of existing approaches. *Th. Comp. Sys.*, 63:715–808, 2019.

[51] A. Gheorghiu and T. Vidick. Computationally-secure and composable remote state preparation. 2019. arXiv:1904.06320.

[52] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory. Robust online Hamiltonian learning. *New J. Phys.*, 14:103013, 2012.

[53] 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.

[54] O. Guehne, M. Reimpell, and R. F. Werner. Estimating entanglement measures in experiments. *Phys. Rev. Lett.*, 98:110502, 2007.

[55] O. Guehne and G. Toth. Entanglement detection. *Phys. Rep.*, 474:1, 2009.

[56] M. Guta, J. Kahn, R. Kueng, and J. A. Tropp. Fast state tomography with optimal error bounds. 2018. arXiv:1809.11162.

[57] M. Guta, T. Kypraios, and I. Dryden. Rank-based model selection for multiple ions quantum tomography. *New J. Phys.*, 14:105002, 2012.

[58] D. Hangleiter, M. Kliesch, J. Eisert, and C. Gogolin. Sample complexity of device-independently certified" quantum supremacy". *Phys. Rev. Lett.*, 122:210502, 2019.

[59] D. Hangleiter, M. Kliesch, M. Schwarz, and J. Eisert. Direct certification of a class of quantum simulations. *Quantum Sci. Technol.*, 2:015004, 2017.

[60] R. Harper, I. Hincks, C. Ferrie, S. T. Flammia, and J. J. Wallman. Statistical analysis of randomized benchmarking. *Phys. Rev. A*, 99:052350, 2019.

[61] J. Helsen, J. J. Wallman, S. T. Flammia, and S. Wehner. Multiqubit randomized benchmarking using few samples. *Phys. Rev. A*, 100(3):032304, 2019.

[62] J. Helsen, X. Xue, L. M. K. Vandersypen, and S. Wehner. A new class of efficient randomized benchmarking protocols. *npj Quant. Inf.*, 5:1–9, 2019.

[63] M. Holzäpfel, T. Baumgratz, M. Cramer, and M. B. Plenio. Scalable reconstruction of unitary processes and Hamiltonians. *Phys. Rev. A*, 91:042129, 2015.

[64] Z. Hradil. Quantum-state estimation. *Phys. Rev. A*, 55:1561–1564, 1997.

[65] Z. Hradil, J. Rehacek, J. Fiurasek, and M. Jezek. Three maximum-likelihood methods in quantum mechanics. In *Quantum state estimation*, pages 59–112. Springer, 2004.

[66] R. Hübener, A. Mari, and J. Eisert. Wick's theorem for matrix product states. *Phys. Rev. Lett.*, 110:040401, 2013.

[67] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White. Measurement of qubits. *Phys. Rev. A*, 64:052312, 2001.

[68] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos. Observation of entanglement propagation in a quantum many-body system. *Nature*, 511:202, 2014.

[69] A. Kalev, R. L. Kosut, and I. H Deutsch. Quantum tomography protocols with positivity are compressed sensing protocols. *npj Quant. Inf.*, 1:15018, 2015.

[70] H. J. Kimble. The quantum internet. *Nature*, 453:1023–1030, 2008.

[71] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki. Robust extraction of tomographic information via randomized benchmarking. *Phys. Rev. X*, 4:011050, 2014.

[72] M. Kliesch, R. Kueng, J. Eisert, and D. Gross. Guaranteed recovery of quantum processes from few measurements. *Quantum*, 3:171, 2019.

[73] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, 2008.

[74] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller. Self-verifying variational quantum simulation of the lattice Schwinger model. *Nature*, 569:355, 2019.

[75] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia. Comparing experiments to the fault-tolerance threshold. *Phys. Rev. Lett.*, 117:170502, 2016.

[76] B. Lévi, C. C. López, J. Emerson, and D. G. Cory. Efficient error characterization in quantum information processing. *Phys. Rev. A*, 75:022314, 2007.

[77] N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe. Experimental comparison of two quantum computing architectures. *Proc. Natl. Ac. Sc.*, 114:3305–3310, 2017.

[78] E. Magesan, J. M. Gambetta, and J. Emerson. Robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:042311, 2011.

[79] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. Da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109:080505, 2012.

[80] U. Mahadev. Classical Verification of Quantum Computations. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 259–267, 2018.

[81] A. Mari, K. Kieling, B. Melholt Nielsen, E.S. Polzik, and J. Eisert. Directly estimating non-classicality. *Phys. Rev. Lett.*, 106:010403, 2011.

[82] D. Markham and A. Krause. A simple protocol for certifying graph states and applications in quantum networks. 2018. arXiv:1801.05057.

[83] D. Mayers and A. Yao. Self testing quantum apparatus. *Quantum. Info. Comput.*, 4:273, 2004.

[84] M. McKague. Self-testing graph states. In *Theory of Quantum Computation, Communication, and Cryptography*, Lecture Notes in Computer Science, pages 104–120. Springer, Berlin, Heidelberg, 2011.

[85] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Corcoles, B. R. Johnson, C. A. Ryan, and M. Steffen. Self-consistent quantum process tomography. *Phys. Rev. A*, 87:062119, 2013.

[86] D. Mogilevtsev, J. Rehacek, and Z. Hradil. Self-calibration for self-consistent tomography. *New J. Phys.*, 14:095001, 2012.

[87] L. Motka, B. Stoklasa, J. Rehacek, Z. Hradil, V. Karasek, D. Mogilevtsev, G. Harder, C. Silberhorn, and L. L. Sanchez-Soto. Efficient algorithm for optimizing data-pattern tomography. *Phys. Rev. A*, 89:054102, May 2014.

[88] A. Natarajan and T. Vidick. Low-degree testing for quantum states. 2018. arXiv:1801.03821.

[89] M. Ohliger, V. Nesme, and J. Eisert. Efficient and feasible state tomography of quantum many-body systems. *New J. Phys.*, 15:015024, 2013.

[90] E. Onorati, A. H. Werner, and J. Eisert. Randomized benchmarking for individual quantum gates. *Phys. Rev. Lett.*, 123:060501, 2019.

[91] S. Pallister, N. Linden, and A. Montanaro. Optimal verification of entangled states with local measurements. *Phys. Rev. Lett.*, 120:170502, 2018.

[92] D. S. Phillips, M. Walschaers, J. J. Renema, I. A. Walmsley, N. Treps, and J. Sperling. Benchmarking of Gaussian boson sampling using two-point correlators. *Phys. Rev. A*, 99:023836, 2019.

[93] C. Portmann and R. Renner. Cryptographic security of quantum key distribution. 2014. arXiv:1409.3525.

[94] J. Preskill. Quantum computing and the entanglement frontier. *Bull. Am. Phys. Soc.*, 58, 2013.

[95] T. J. Proctor, A. Carignan-Dugas, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young. Direct randomized benchmarking for multiqubit devices. *Phys. Rev. Lett.*, 123:030503, 2019.

[96] O. Regev. The learning with errors problem. *Invited survey in CCC*, 7, 2010.

[97] J. Rehacek, D. Mogilevtsev, and Z. Hradil. Operational tomography: Fitting of data patterns. *Phys. Rev. Lett.*, 105:010402, 2010.

[98] D. M. Reich, G. Gualdi, and C. P. Koch. Optimal strategies for estimating the average fidelity of quantum gates. *Phys. Rev. Lett.*, 111:200401, 2013-12.

[99] B. W. Reichardt, F. Unger, and U. Vazirani. Classical command of quantum systems. *Nature*, 496:456–460, 2013.

[100] A. Rocchetto. Stabiliser states are efficiently PAC-learnable. *Quant. Inf. Comp.*, (7&8):541–552, 2018.

[101] A. Rocchetto, S. Aaronson, S. Severini, G. Carvacho, D. Poderini, I. Agresti, M. Bentivegna, and F. Sciarrino. Experimental learning of quantum states. *Science Adv.*, 5:eaau1946, 2019.

[102] I. Roth, R. Kueng, S. Kimmel, Y.-K. Liu, D. Gross, J. Eisert, and M. Kliesch. Recovering quantum gates from few average gate fidelities. *Phys. Rev. Lett.*, 121:170502, 2018.

[103] V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev. The security of practical quantum key distribution. *Rev. Mod. Phys.*, 81(3):1301–1350, September 2009.

[104] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch. Observation of many-body localization of interacting fermions in a quasi-random optical lattice. *Science*, 349:842–845, 2015.

[105] B. Schumacher. Sending entanglement through noisy quantum channels. *Phys. Rev. A*, 54:2614–2628, 1996.

[106] P. Sekatski, J.-D. Bancal, S. Wagner, and N. Sangouard. Certifying the building blocks of quantum computers from Bell's theorem. *Phys. Rev. Lett.*, 121:180505, 2018.

[107] J. Y. Sim, J. Shang, H. K. Ng, and B.-G. Englert. Proper error bars for self-calibrating quantum tomography. 2019. arXiv: 1904.11202.

[108] N. Spagnolo, C. Vitelli, M. Bentivegna, D. J. Brod, A. Crespi, F. Flamini, S. Giacomini, G. Milani, R. Ramponi, P. Mataloni, R. Osellame, E. F. Galvao, and F. Sciarrino. Efficient experimental validation of photonic boson sampling against the uniform distribution. *Nature Phot.*, 8:615–620, 2014.

[109] A. Streltsov, G. Adesso, and M. B. Plenio. Colloquium: Quantum coherence as a resource. *Rev. Mod. Phys.*, 89:041003, 2017.

[110] I. Supic and J. Bowles. Self-testing of quantum systems: a review. 2019. arXiv:1904.10042.

[111] Y. Takeuchi, A. Mantri, T. Morimae, A. Mizutani, and J. F. Fitzsimons. Resource-efficient verification of quantum computing using Serflings bound. *npj Quantum Inf*, 5:1–8, 2019.

[112] Y. Takeuchi and T. Morimae. Verification of many-qubit states. *Phys. Rev. X*, 8:021060, 2018.

[113] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo. Many-body quantum state tomography with neural networks. *Nature Physics*, 14:447–450, 2018.

[114] S. Trotzky, Y-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch. Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas. *Nature Physics*, 8:325–330, 2012.

[115] S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N.V. Prokof'ev, B. Svistunov, and M. Troyer. Suppression of the critical temperature for superfluidity near the mott transition: validating a quantum simulator. *Nature Phys.*, 6:998–1004, 2010.

[116] G. Valiant and P. Valiant. An automatic inequality prover and instance optimal identity testing. *SIAM J. Comput.*, 46:429–455, 2017.

[117] J. Wallman, C. Granade, R. Harper, and S. T. Flammia. Estimating the coherence of noise. *New J. Phys.*, 17:113020, 2015.

[118] J. J. Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2:47, 2018.

[119] J. J. Wallman, M. Barnhill, and J. Emerson. Robust characterization of loss rates. *Phys. Rev. Lett.*, 115:060501, 2015.

[120] J. J. Wallman, M. Barnhill, and J. Emerson. Robust characterization of leakage errors. *New J. Phys.*, 18:043021, 2016.

[121] J. J. Wallman and S. T. Flammia. Randomized benchmarking with confidence. *New. J. Phys.*, 16:103032, 2014.

[122] J. Wang, V. B. Scholz, and R. Renner. Confidence polytopes in quantum state tomography. *Phys. Rev. Lett.*, 122:190401, 2019.

[123] J. Watrous. Simpler semidefinite programs for completely bounded norms. 2012. arXiv:1207.5726.

[124] S. Wehner, D. Elkouss, and R. Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412).