

Overview of projective quantum measurements

Diego Barberena^{1,2} and Aaron J. Friedman^{1,*}

¹*Department of Physics and Center for Theory of Quantum Matter,
University of Colorado, Boulder CO 80309, USA*

²*JILA and NIST, University of Colorado, Boulder, CO 80309, USA*

(Dated: April 10, 2024)

We provide an overview of standard “projective” quantum measurements with the goal of elucidating connections between theory and experiment. We make use of a unitary “Stinespring” representation of measurements on a dilated Hilbert space that includes both the physical degrees of freedom and those of the measurement apparatus. We explain how this unitary representation (*i*) is guaranteed by the axioms of quantum mechanics, (*ii*) relates to both the Kraus and von Neumann representations, and (*iii*) corresponds to the physical time evolution of the system and apparatus during the measurement process. The Stinespring representation also offers significant conceptual insight into measurements, helps connects theory and experiment, is particularly useful in describing protocols involving midcircuit measurements and outcome-dependent operations, and establishes that all quantum operations are compatible with relativistic locality, among other insights.

CONTENTS

1. Introduction	2
2. Mathematical formalism for measurements	4
2.1. Measurement and spectral decomposition of observables	5
2.2. Kraus representation	6
2.3. Unitary representation	8
2.4. von Neumann representation	12
2.5. Summary of measurement representations	15
3. Photon measurements	17
3.1. Electromagnetic modes	17
3.2. Destructive measurement of photon number	18
3.3. Balanced homodyne detection	21
4. Qubit measurements	24
4.1. Fluorescence measurement	24
4.2. Dispersive measurement	28
5. Using the unitary measurement formalism	31
5.1. Measurement outcomes and the Born rule	31
5.2. Application to measurement-based protocols	32
5.3. Decoherence and “wavefunction collapse”	34
5.4. Absence of “spooky action at a distance”	38
5.5. Quantum mechanics is local	40
6. Conclusion	42

* Aaron.Friedman@colorado.edu

A. Details of the Stern-Gerlach experiment	44
A.1. Time evolution of operators in the general case	44
A.2. Time evolution of states in the simple case	45
B. Photon absorption	46
C. Matrix element for homodyne detection	48
References	50

1. INTRODUCTION

Measurements are a fundamental part of quantum theory. Typically codified in axiomatic formulations of quantum mechanics [1–8], the standard “projective” measurement is intrinsically related to the most prominent departures of quantum systems from more familiar classical physics. For example, the quantization of observable outcomes and the probabilistic nature of measurements gave some of the first historical indications of the inadequacy of classical mechanics in describing, e.g., atomic systems. Despite the successes of quantum mechanics, the nature of measurements—and their implications for quantum mechanics as a theory—has been a source of substantial debate.

The apparent “collapse” of the wavefunction upon measurement was one of several early sources of confusion and concern. The probabilistic nature of measurements was seen as a problem to those who expected a deterministic theory, à la classical physics [9]. Moreover, the putative violation of relativistic locality in the measurement of spatially separated entangled degrees of freedom was viewed as a “paradox” that implied that quantum theory was not yet complete [10]. While it is now understood that no such violation exists, measurements remain one of the more “mysterious” aspects of quantum mechanics, and are at the heart of debates over the theory’s interpretation.

We stress that addressing such foundational questions (e.g., about interpretations of quantum mechanics) is *not* our goal. However, doing so would require considering the details of how measurements are implemented in real experiments. Importantly, we note that the standard theoretical descriptions of quantum measurements [11–13] obfuscate their nature. Measurements and other nonunitary quantum operations are typically modeled via (i) stochastic updates to the wavefunction, (ii) a Lindblad master equation [14–16], or (iii) Kraus operators [17–20], all of which describe only the degrees of freedom in the system of interest, despite the fact that all such quantum operations are exclusive to *open* systems. Even setting aside matters of interpretation, consideration of the experimental details surrounding measurement is important for connecting theory to experiment and designing measurement-based protocols for near-term quantum devices.

To that end, we consider a conceptually transparent and analytically powerful representation of projective measurements in terms of a unitary operator $\mathcal{U}$ acting on an enlarged Hilbert space [21, 22]. This representation is logically implied by all axiomatic formulations of quantum mechanics [1–8], and describes precisely the unitary time evolution of the physical system and measurement apparatus in the experimental implementation of the prototypical examples of measurements considered herein. Although an exhaustive study of all possible implementations of projective measurements is beyond the scope of this work, we expect that $\mathcal{U}$ always has this physical interpretation. In addition to elucidating various aspects of measurements—e.g., related to their experimental implementation and compatibility with both relativistic locality and determinism—the unitary representation we showcase is also better suited to the investigation of measurement-based protocols involving projective measurements and outcome-dependent operations. We emphasize this because such protocols have immense utility in numerous tasks relevant to near-term quantum technologies, from quantum information processing to efficient many-body state preparation [21–25].

The primary goals of this paper are to (i) provide a useful overview of projective quantum measurements in theory and experiment; (ii) explain how these measurements are most commonly implemented in practice; (iii) explain the rigorous mathematical origins of the unitary “Stinespring” representation of projective measurements; (iv) establish how the unitary representation corresponds to the time evolution of the system and detector during the measurement process; and (v) explain how to apply the Stinespring formalism to generic quantum protocols, its benefits (e.g., in the context of protocols with midcircuit measurements and outcome-dependent feedback), and some of its more important implications. The remainder of this paper is organized as follows.

In Sec. 2 we derive the unitary measurement formalism from the axioms of quantum mechanics [1–8], using powerful results from C^* -algebras [26–31]. In Sec. 2.1, we discuss the axiomatic properties of quantum measurements. In Sec. 2.2, we derive the standard *Kraus representation* of measurements [17–20] using both the Stinespring dilation theorem [28] and Choi’s theorem [29] for completely positive (CP) maps. In Sec. 2.3.1, we explain the “physicist’s Stinespring theorem” [30, 31], in which quantum operations are captured by a *unitary* operator $\mathcal{U}$ on a *dilated* Hilbert space $\mathcal{H}_{\text{dil}}$ (2.8), which includes both the physical Hilbert space $\mathcal{H}$ and a Stinespring Hilbert space $\mathcal{H}_{\text{ss}}$. This result does not follow straightforwardly from Stinespring’s theorem [28], but was proven by Kraus [20] for finite-dimensional Hilbert spaces; a proof for the infinite-dimensional case $\mathcal{H} = L^2(\mathbb{R}) = \ell^2(\mathbb{N}) \cong \mathbb{C}^\infty$ is the subject of forthcoming work [31]. In Sec. 2.4.1, we connect $\mathcal{U}$ to von Neumann’s description of measurements using a *pointer Hamiltonian* [32, 33] that includes the detector. All of the above is summarized in Sec. 2.5

In Secs. 3 and 4, we discuss the experimental implementations of measurements of photons and qubits, respectively. While quantum systems can be probed in a variety of ways, the final stage of these probes is typically selected from a small set of measurements that includes the examples in Secs. 3 and 4. We also highlight that most measurement schemes involve electromagnetic modes (e.g., photons), and typically record outcomes through particle detection, which generally involve the excitation of particles in an apparatus. In Sec. 3.1, we review the quantum description of electromagnetic fields. In Sec. 3.2, we consider measurements of photon number (or intensity); despite being “destructive” (rather than projective), these measurements are captured by the unitary representation. In Sec. 3.3, we discuss interferometry measurements involving homodyne detection, which culminate in intensity measurements. In Sec. 4.1, we review fluorescence measurements of various types of qubits, which also culminate in counting spontaneously emitted photons. In Sec. 4.2, we discuss “dispersive readout” of the states of, e.g., superconducting qubits.

The examples in Secs. 3 and 4 illustrate that the dilated Hilbert space $\mathcal{H}_{\text{dil}}$ (2.8) is *physical*. We note that the *minimal* Stinespring Hilbert space $\mathcal{H}_{\text{ss}}$ that we derive in Sec. 2 has a basis $\{|m\rangle_{\text{ss}}\}$ corresponding to the $\mathcal{N}$ possible measurement outcomes. However, in the examples considered in Secs. 3 and 4, we generally identify a nonminimal dilated unitary representation $\mathcal{U}$ involving additional degrees of freedom and/or multiple states reflecting the same outcome. Importantly, appropriately “binning” the states of $\mathcal{H}_{\text{ss}}$ corresponding to the same outcome leads to the minimal Stinespring representation, which is unique up to the choice of “default” initial state of the apparatus.

Finally, in Sec. 5, we discuss how to use the unitary formalism, and some of its implications. In Sec. 5.1, we show how the usual Born rule comes from and extract statistics. In Sec. 5.2, we explain the utility of the dilated unitary representation $\mathcal{U}$ in describing measurement-based protocols. Because the outcomes are stored in explicit degrees of freedom, it is far more straightforward in the unitary representation to describe quantum operations conditioned on the outcomes of prior measurements than it is using the Kraus representation. We emphasize that, given the utility of measurement-based protocols, this is one of the main advantages of the unitary formalism. In Sec. 5.3, we explain the appearance of “wavefunction collapse” using a simple model of decoherence [15] based on recent progress on many-body quantum chaos—the mechanism underlying thermalization. Although the dilated unitary representation $\mathcal{U}$ is fully *deterministic*,

with all outcomes occurring, the fact that the measurement apparatus is classical (so that the outcome can be read off) also ensures that it is subject to decoherence [15]. However, because the outcomes are associated with a symmetry charge (e.g., particle number), superpositions of distinct outcomes rapidly decohere into a mixed state, so that only one outcome is observed. In Sec. 5.4, we explain the absence of “spooky action at a distance” in measurements of entangled states, as (i) no information is transferred and (ii) the intrinsic nature of collapse and absence of “branching” of the dilated wavefunction upon measurement means there is no “influence” of one particle in an entangled pair on the other. Lastly, in Sec. 5.5, we explain how the unitary representation establishes that *all quantum operations involving measurements obey relativistic locality* [21].

2. MATHEMATICAL FORMALISM FOR MEASUREMENTS

We begin with a careful derivation of mathematical representations of projective and similar measurements, culminating in the unitary “Stinespring” formulation. The Stinespring unitary $\mathcal{U}$ connects directly to the Kraus [17–20] and von Neumann [2, 32] representations of measurements, and acts on a dilated Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H} \otimes \mathcal{H}_{\text{ss}}$ (2.8). While the degrees of freedom in $\mathcal{H}_{\text{ss}}$ are nominally just a bookkeeping tool, the examples in Secs. 3 and 4 show that they are *physical*, corresponding to the measurement apparatus. Readers for whom such a unitary formulation of measurements is familiar and/or intuitive may prefer to skip to the summary in Sec. 2.5.

The unitary representation of measurements we discuss below follows from mathematical results for C^* -algebras [20, 26–29]. Historically, C^* -algebras were developed to describe the algebraic properties of the operators associated with quantum systems. In fact, every axiomatic formulation of quantum mechanics implies a C^* -algebra corresponding to the “bounded” operators $\mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ [1–8, 26–31]; moreover, the allowed updates to a quantum system—known as *quantum operations*—correspond to completely positive (CP) maps between C^* -algebras [20, 24].

In particular, the density matrix ρ describing a quantum system always belongs to $\mathcal{B}(\mathcal{H})$. Density matrices are positive operators $\rho \geq 0$, which means that their spectra (i.e., eigenvalues) are real and nonnegative. They also have unit trace $\text{tr}(\rho) = 1$, so their eigenvalues encode a probability distribution. Because a quantum operation Φ is a CP map, it sends positive operators ρ to positive operators $\Phi(\rho) \geq 0$, even if ρ only describes part of a larger system. An important subset of quantum operations are known as *quantum channels*, which are both completely positive and trace preserving (CPTP). When Φ is a quantum channel, $\rho' = \Phi(\rho)$ is also a density matrix. However, generic quantum operations are *trace decreasing*, meaning that $\Phi(\rho) = \lambda \rho'$ for some density matrix ρ' and real number $0 \leq \lambda \leq 1$. We consider quantum operations to allow for, e.g., measurements resulting in a particular outcome m , for which $\Phi_m(\rho) = p_m \rho_m$, with p_m the probability of outcome m and ρ_m the corresponding (and normalized) postmeasurement density matrix. Note that the map $\Phi_m : \rho \mapsto p_m \rho_m$ is always well defined and linear in ρ , while a map $\rho \mapsto \rho_m$ is ill defined when $p_m = 0$, and is always a nonlinear function of ρ , since p_m depends on ρ [20, 24].

The correspondence between quantum operations and CP maps between C^* -algebras provides for the derivations below [30, 31], due to useful properties of C^* -algebras and maps [20–22, 26–31]. In particular, the Stinespring dilation theorem [28] implies a representation of any quantum operation Φ on a *dilated* Hilbert space $\mathcal{H}_{\text{dil}}$. In the case of finite-dimensional Hilbert spaces, Choi’s theorem [29] implies a representation of Φ in terms of *Kraus operators* [17–20], which is the most common representation of measurements in the literature. However, the Kraus formulation also implies the *unitary* representation of measurements on $\mathcal{H}_{\text{dil}}$ that we explore herein. The unitary “Stinespring” representation results from a “physicist’s version” of the Stinespring dilation theorem [28] discussed in Sec. 2.3.1, which is actually due to Kraus [20]. We refer to the unitary formalism as the “Stinespring representation” due to its connection to the physicist’s Stinespring theorem and

to avoid confusion with the Kraus representation in terms of Kraus operators. A rigorous proof for the infinite-dimensional Hilbert space $\mathcal{H} = \ell^2(\mathbb{N})$ is the subject of forthcoming work [31].

In Sec. 2.1 we discuss the axiomatic properties of quantum measurements. In Sec. 2.2, we derive the Kraus representation [17–20] from Stinespring’s dilation theorem [28] and Choi’s theorem [29]. In Sec. 2.3, we derive the *unitary* representation of measurements on a $\mathcal{H}_{\text{dil}}$, which we relate to the Kraus formulation. In Sec. 2.4, we discuss von Neumann’s “pointer” Hamiltonian describing measurements [32], which we relate to the unitary representation. Finally, for convenience, we summarize the foregoing results for measurement representations in Sec. 2.5.

Although we regard the extra Stinespring degrees of freedom $\mathcal{H}_{\text{ss}} \subset \mathcal{H}_{\text{dil}}$ as a bookkeeping tool below, in Secs. 3 and 4 we show that $\mathcal{H}_{\text{ss}}$ is *physical* in several prominent examples of measurements, and we expect that it holds for generic projective and destructive measurements. In particular, $\mathcal{H}_{\text{ss}}$ reflects the state space of the *measurement apparatus*, where the dilated measurement unitary $\mathcal{U}$ (2.17)—or, equivalently, the von Neumann Hamiltonian H_{vN} (2.34)—captures time evolution of the system and apparatus during the measurement process. Again, we expect this unitary description of measurements to be intuitive to many readers, who may prefer to skip to Sec. 2.5.

2.1. Measurement and spectral decomposition of observables

Consider the standard, projective measurement of a generic observable $\mathcal{O}$ in a system described by the density matrix ρ . Suppose that $\mathcal{O}$ has $\mathcal{N}$ *unique* eigenvalues $\{O_m\}$ for $0 \leq m < \mathcal{N}$, and denote by r_m the multiplicity (degeneracy) of the m th eigenvalue O_m . We then identify a set of *eigenprojectors* onto the $\mathcal{N}$ distinct eigenspaces of $\mathcal{O}$, i.e.,

$$\mathcal{O} P_m = O_m P_m \quad \text{with} \quad \text{tr}(P_m) = r_m, \quad (2.1)$$

where r_m is also the “rank” of P_m , meaning that $P_m = |m\rangle\langle m|$ when $r_m = 1$, and otherwise, $P_m = \sum_{k=1}^{r_m} |m_k\rangle\langle m_k|$ is a sum over projectors onto degenerate eigenstates $|m_k\rangle$ of $\mathcal{O}$ with eigenvalue O_m . The eigenstates are orthonormal and form a complete basis for $\mathcal{H}$ [1–8]. The $\mathcal{N}$ projectors (2.1) form an orthogonal and complete set, meaning that,

$$\sum_{m=0}^{\mathcal{N}-1} P_m = \mathbb{1} \quad \text{and} \quad P_m P_n = \delta_{m,n} P_m, \quad (2.2)$$

and using these projectors, we define the *spectral decomposition* of the observable $\mathcal{O}$ via

$$\mathcal{O} = \sum_{m=0}^{\mathcal{N}-1} O_m P_m, \quad (2.3)$$

which is unique, and holds even in the limit $\mathcal{N} \rightarrow \infty$.

The axioms of quantum mechanics [1–8] dictate that the outcome of projectively measuring $\mathcal{O}$ (2.3) is one of its eigenvalues O_n . Repeating the experiment many times, the expectation value (i.e., average) of the measurement of $\mathcal{O}$ in the state ρ is given by $\langle \mathcal{O} \rangle_\rho = \text{tr}(\mathcal{O} \rho)$. However, the probability-theoretic definition of the expectation value also implies that

$$\langle \mathcal{O} \rangle_\rho = \sum_{m=0}^{\mathcal{N}-1} O_m p_m = \text{tr}(\mathcal{O} \rho) = \sum_{m=0}^{\mathcal{N}-1} O_m \text{tr}(P_m \rho) \implies p_m = \text{tr}(P_m \rho), \quad (2.4)$$

where p_m is the probability of observing outcome O_m upon measuring $\mathcal{O}$ in the state ρ . When $\rho = |\psi\rangle\langle\psi|$ is a pure state and $r_m = 1$, then $p_m = \langle m|\psi\rangle^2$, reproducing the familiar Born rule.

If $\mathcal{O}$ (2.3) is measured in the state ρ and the observed outcome is O_m , then the postmeasurement state *collapses* into a state ρ_m [1–8]; up to normalization, we have that

$$\rho \mapsto \rho_m \propto P_m \rho P_m \quad \text{where} \quad \text{tr}(P_m \rho P_m) = \text{tr}(P_m \rho) = p_m, \quad (2.5)$$

meaning that the normalized density matrix ρ_m can only be defined if $p_m > 0$. Accordingly, we identify a quantum operation (or linear CP map) Φ_m , corresponding to a measurement of $\mathcal{O}$ in the state ρ resulting in the particular outcome O_m , which takes the form

$$\Phi_m(\rho) = P_m \rho P_m = p_m \rho_m, \quad (2.6)$$

where $\rho_m = 0$ if $p_m = 0$, and $\rho_m = P_m \rho P_m / p_m$ otherwise. Note that if $\mathcal{O}$ is measured again in the postmeasurement state ρ_m , the probability to obtain outcome O_m again is unity, as expected.

We also separately define a quantum *channel* corresponding to the measurement of $\mathcal{O}$, given by

$$\Phi(\rho) \equiv \sum_{m=1}^{\mathcal{N}} \Phi_m(\rho) = \sum_{m=1}^{\mathcal{N}} P_m \rho P_m = \sum_{m=1}^{\mathcal{N}} p_m \rho_m \equiv \rho_{\text{av}}, \quad (2.7)$$

which is CP, linear, and trace preserving, since $\text{tr}(\rho_{\text{av}}) = \sum_{m=1}^{\mathcal{N}} \text{tr}[\Phi_m(\rho)] = \sum_{m=1}^{\mathcal{N}} p_m = 1$. We comment that nonprojective measurements (e.g., weak and generalized measurements) need not satisfy the foregoing properties, and we relegate their consideration to future work. As a reminder, we expect that all of the statements above hold true even in the countably infinite case $\mathcal{N} \rightarrow \infty$.

2.2. Kraus representation

The starting point for all representations of quantum operations is the Stinespring dilation theorem [28] for CP maps between C^* -algebras [26, 27]. As a reminder, all density matrices ρ belong to the C^* -algebra $\mathcal{B}(\mathcal{H})$, corresponding to the bounded operators on a Hilbert space $\mathcal{H}$; the quantum operations of interest are CP maps from $\mathcal{B}(\mathcal{H})$ to itself. The Stinespring dilation theorem [28] establishes that all CP maps can be represented on a *dilated* Hilbert space

$$\mathcal{H}_{\text{dil}} = \mathcal{H} \otimes \mathcal{H}_{\text{ss}}, \quad (2.8)$$

using the following combination of completely positive operations,

$$\Phi(\rho) = V^\dagger \pi(\rho) V, \quad (2.9)$$

where $V : \mathcal{H} \rightarrow \mathcal{H}_{\text{dil}}$ is an *isometry*—i.e., a length-preserving CP map—and $\pi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H}_{\text{dil}})$ is a $*$ -homomorphism—i.e., a structure-preserving CP map between C^* -algebras, whose properties are unimportant to the present discussion. Further simplification is not possible in full generality.

However, when $\mathcal{H}$ is finite dimensional [30, 31], useful simplifications *are* possible. We also expect that this extends to the countably infinite case $\mathcal{H} = \ell^2(\mathbb{N})$ via inductive limits [31, 34–36]. For finite-dimensional $\mathcal{H}$, we have that $\mathcal{B}(\mathcal{H}) \cong M_N(\mathbb{C})$, where $M_N(\mathbb{C})$ is the C^* -algebra of $N \times N$ matrices with complex entries, for which it is well known [26, 27, 30, 31] that $\pi(\rho)$ in Eq. 2.9 can be written as $\pi(\rho) = u^\dagger(\rho \otimes \mathbb{1}_{\text{ss}})u$ for some unitary u acting on $\mathcal{H}_{\text{dil}}$. Accordingly, we have that

$$\Phi(\rho) = V^\dagger \rho \otimes \mathbb{1}_{\text{ss}} V, \quad (2.10)$$

where we absorb u into V without loss of generality. We note that the representation above also invokes Choi’s theorem for CP maps between finite-dimensional C^* -algebras [29].

For CP maps between generic C^* -algebras, which may be infinite dimensional, we stress that Eq. 2.10 is not guaranteed to hold. This is because the $*$ -homomorphism π in Eq. 2.9 is not guaranteed to have a convenient functional form beyond finite-dimensional C^* -algebras, which are isomorphic to $M_N(\mathbb{C})$. However, if $\mathcal{H}$ is a tensor product of finitely many countably infinite Hilbert spaces $\mathcal{H}_i = \ell^2(\mathbb{N})$, we expect that the Stinespring representation (2.10) remains valid [31], as these Hilbert spaces can be realized as inductive limits of a sequence of finite-dimensional Hilbert spaces $\mathcal{H}$. The inductive limit of $M_N(\mathbb{C})$ is the set $\mathcal{K}(\mathcal{H})$ of *compact* operators on $\mathcal{H} = \ell^2(\mathbb{N})$ —an *approximately finite* C^* -algebra [31, 34–36] to which all density matrices ρ belong. Moreover, any observable $\mathcal{O} \in \text{End}(\mathcal{H})$ as $\dim(\mathcal{H}) \rightarrow \infty$ is guaranteed to be close in the weak operator topology to an element of the approximately finite C^* -algebra $\mathcal{K}(\mathcal{H})$ [26, 27, 31, 34–36].

We now derive the *Kraus representation* of Φ . Suppose that $\mathcal{H}_{\text{ss}}$ has dimension $\mathcal{D}_{\text{ss}} = q$ and an orthonormal basis $\{|k\rangle\}$; resolving the identity in Eq. 2.10, we find that

$$\Phi(\rho) = \sum_{k=1}^q V^\dagger \rho \otimes |k\rangle\langle k|_{\text{ss}} V = \sum_{k=1}^q E_k \rho E_k^\dagger \quad \text{with} \quad E_k \equiv V^\dagger |k\rangle_{\text{ss}}, \quad (2.11)$$

where $E_k : \mathcal{H} \rightarrow \mathcal{H}$ is a Kraus operator on $\mathcal{H}$ and $q \leq \mathcal{D}^2$ (when $\mathcal{D}$ is finite). When the map Φ (2.11) is a quantum channel (i.e., trace preserving), then the Kraus operators form a *complete set*,

$$0 = \text{tr}[\rho - \Phi(\rho)] = \text{tr} \left[\rho \left(\mathbb{1} - \sum_{k=1}^q E_k^\dagger E_k \right) \right] \implies \sum_{k=1}^q E_k^\dagger E_k = \mathbb{1}, \quad (2.12)$$

much like the projectors P_m (2.1) form a complete set (2.2). When the map Φ is a quantum operation (i.e., trace decreasing), we instead find that the Kraus operators form an *incomplete set*,

$$0 \leq \text{tr}[\rho - \Phi(\rho)] = \text{tr} \left[\rho \left(\mathbb{1} - \sum_{k=1}^q E_k^\dagger E_k \right) \right] \implies E_0^2 \equiv \mathbb{1} - \sum_{k=1}^q E_k^\dagger E_k \geq 0, \quad (2.13)$$

where the new Kraus operator E_0 on the right is well defined because it is the square root of a positive operator (i.e., its eigenvalues are real and positive). Importantly, this implies that *every* incomplete set of Kraus operators corresponding to a quantum operation is part of a larger, complete set of Kraus operators, since one can always append E_0 to the incomplete set to recover

$$\sum_{k=0}^q E_k^\dagger E_k = \mathbb{1}, \quad (2.14)$$

meaning that every trace-decreasing quantum operation Φ involves a subset of a larger, complete Kraus operators that define a related, trace-preserving quantum channel.

Lastly, we discuss the Kraus representation of *measurements*. First, consider the trace-decreasing quantum operation Φ_m corresponding to a measurement of $\mathcal{O}$ (2.3) resulting in the m th outcome O_m , so that $\Phi_m(\rho) = p_m \rho_m$ (2.6) where $p_m = \text{tr}(\rho P_m)$ is the probability to recover outcome m (2.4) and ρ_m is the collapsed postmeasurement density matrix following observation of outcome m . The Kraus representation (2.11) that realizes the postmeasurement state $\Phi_m(\rho) = p_m \rho_m$ (2.6) involves a single Kraus operator $E_m = P_m$ (2.1), with $q = 1$. Next, consider the trace-preserving quantum channel Φ_{av} corresponding to the average over a large number of measurements of $\mathcal{O}$ (2.3), meaning that $\Phi_{\text{av}}(\rho) = \rho_{\text{av}} = \sum_m p_m \rho_m = \sum_m \Phi_m(\rho)$ (2.7). There are $q = \mathcal{N}$ Kraus operators, given by $E_m = P_m$ (2.1), corresponding to the $\mathcal{N}$ possible outcomes (the unique eigenvalues of $\mathcal{O}$). Note that the set $\{P_m\}$ (2.1) of Kraus operators is complete (2.2), as required (2.14).

2.3. Unitary representation

We now use the Kraus representation (2.11) to derive a *unitary* representation of quantum operations and channels Φ on the dilated Hilbert space $\mathcal{H}_{\text{dil}}$. This representation (2.15) is sometimes misattributed to the Stinespring dilation theorem [26–28] in the literature, though it was actually proven by Kraus [20] for finite-dimensional quantum systems. To avoid confusion with the Kraus representation discussed in Sec. 2.2, we refer to Kraus’s unitary representation of CP maps Φ as the “Stinespring representation,” and the corresponding theorem as the “physicist’s Stinespring theorem.” We expect this proof to extend—using the machinery of approximately finite C^* -algebras [34–36] and inductive and projective limits—to the infinite-dimensional Hilbert space $\ell^2(\mathbb{N})$ [31].

2.3.1. The physicist’s Stinespring theorem

The physicist’s Stinespring dilation theorem [28, 31] implies that any quantum channel (i.e., CPTP map) Φ acting on a finite-dimensional Hilbert space $\mathcal{H}$ can be written in the form

$$\Phi(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger \right), \quad (2.15)$$

where $\mathcal{U}$ acts unitarily on $\mathcal{H}_{\text{dil}}$ (2.8), $|i\rangle_{\text{ss}} \in \mathcal{H}_{\text{ss}}$ is some “default” initial state of the Stinespring Hilbert space, and the partial trace is over $\mathcal{H}_{\text{ss}}$. Importantly, there is always a *minimal* Stinespring representation—corresponding to the smallest possible dimension of $\mathcal{H}_{\text{ss}}$ [26–28, 30]—which is *unique*, up to the definition of $|i\rangle_{\text{ss}}$. We often consider nonminimal Stinespring representations.

We next extend the unitary representation of channels (2.15) to quantum operations (i.e., trace-decreasing CP maps) Φ_m . The Kraus operators that define a quantum operation Φ_m are a subset of those that define an associated quantum channel Φ (2.15), identified by the $q_m < q = \mathcal{D}_{\text{ss}}$ basis states $\{|k_1\rangle, \dots, |k_{q_m}\rangle\}$ of $\mathcal{H}_{\text{ss}}$ (2.11). Accordingly, the quantum operation Φ_m can be written as

$$\Phi_m(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger P_{\text{ss}}^{(m)} \right), \quad (2.16)$$

where $P_{\text{ss}}^{(m)}$ projects onto the appropriate q_m basis states of $\mathcal{H}_{\text{ss}}$ associated with Φ_m , and $\mathcal{U}$ is the *same* dilated unitary that defines the corresponding quantum channel Φ (2.15) [30, 31].

The dilated unitary $\mathcal{U}$ in Eqs. 2.15 and 2.16 relates to both the Stinespring (2.10) and Kraus (2.11) representations in that it acts on an arbitrary physical state $|\psi\rangle \in \mathcal{H}$ via

$$\mathcal{U}|\psi\rangle \otimes |i\rangle_{\text{ss}} = \sum_{k=1}^q (E_k|\psi\rangle) \otimes |k\rangle_{\text{ss}} = \sum_{k=1}^q \left(V^\dagger |\psi\rangle \otimes |k\rangle_{\text{ss}} \right) \otimes |k\rangle_{\text{ss}}, \quad (2.17)$$

where $|i\rangle_{\text{ss}}$ is the “default” initial state on $\mathcal{H}_{\text{ss}}$, and it is straightforward to check that this operation is length preserving (i.e., isometric) and surjective (i.e., unitary). As with the Kraus representation (2.11) of Sec. 2.2, we expect that the results above extend to infinite-dimensional Hilbert spaces upon taking an inductive limit of finite-dimensional Hilbert spaces [31, 34–36]. We now consider the particular form of this unitary in the context of projective (and similar) measurements.

2.3.2. Finite spectra

We first work out the dilated unitary $\mathcal{U}$ (2.17) corresponding to the projective measurement of an observable $\mathcal{O}$ (2.3) with a finite number $\mathcal{N}$ of unique eigenvalues. The minimum number of Kraus operators (2.11) needed to represent a measurement channel is $q = \mathcal{N}$. For convenience, we

label the eigenvalues of $\mathcal{O}$ by $0 \leq m < \mathcal{N}$, so that the Kraus operators are given by $E_m = P_m$ (2.1). These Kraus operators are associated with a *basis* $\{|m\rangle\}$ of $\mathcal{H}_{\text{ss}}$, and without loss of generality, we initialize the Stinespring register in the default state $|i\rangle_{\text{ss}} = |0\rangle_{\text{ss}}$, so that Eq. 2.17 becomes

$$\mathcal{U} |\psi\rangle \otimes |0\rangle_{\text{ss}} = \sum_{m=0}^{\mathcal{N}-1} P_m |\psi\rangle \otimes |m\rangle_{\text{ss}}, \quad (2.18)$$

which, in the minimal case with $\dim(\mathcal{H}_{\text{ss}}) = \mathcal{N}$, is *uniquely* fulfilled by the dilated unitary operator

$$\mathcal{U}_{[\mathcal{O}]} \equiv \sum_{m=0}^{\mathcal{N}-1} P_m \otimes X_{\text{ss}}^m, \quad (2.19)$$

up to the choice (and definition) of the default state $|0\rangle_{\text{ss}}$. The expression above involves the unitary $\mathcal{N}$ -state Weyl *shift operator* X [22], which acts on $\mathcal{H}_{\text{ss}}$ as

$$X_{\text{ss}}^m \equiv \sum_{k=0}^{\mathcal{N}-1} |k+m \bmod \mathcal{N}\rangle \langle k|_{\text{ss}}, \quad (2.20)$$

and maps the default Stinespring state $|0\rangle_{\text{ss}}$ to $|m\rangle_{\text{ss}}$, as required. The shift operator X is a unitary extension of the Pauli operator X to Hilbert spaces with $\mathcal{N} \geq 2$ [22].

It is straightforward to verify that the Kraus operators $E_m = P_m$ form a complete set (2.14) and that $\mathcal{U}$ (2.19) is unitary on $\mathcal{H}_{\text{dil}}$. We next confirm that the quantum channel Φ (2.15) corresponding to $\mathcal{U}$ (2.19) leads to the outcome-averaged density matrix ρ_{av} (2.7),

$$\Phi(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |0\rangle \langle 0| \mathcal{U}^\dagger \right) = \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho P_n \text{tr}_{\text{ss}} (X^{-m} |0\rangle \langle 0| X^n) = \sum_{m=0}^{\mathcal{N}-1} P_m \rho P_m = \rho_{\text{av}}, \quad (2.21)$$

as required. We also confirm that the quantum operation Φ_ℓ (2.16)—corresponding to a measurement of $\mathcal{O}$ resulting in the outcome O_ℓ —leads to p_ℓ times the collapsed density matrix ρ_ℓ (2.6),

$$\Phi_\ell(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |0\rangle \langle 0| \mathcal{U}^\dagger P_{\text{ss}}^{(\ell)} \right) = \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho P_n \langle \ell | m \rangle \langle n | \ell \rangle = P_\ell \rho P_\ell = p_\ell \rho_\ell, \quad (2.22)$$

as required. In other words, the quantum operation corresponding to a measurement resulting in the outcome ℓ recovers from projecting $\mathcal{H}_{\text{ss}}$ onto its ℓ th basis state.

Thus far, the dilated unitary $\mathcal{U}$ (2.18) has been a bookkeeping tool. In Secs. 3 and 4, we demonstrate how, in real experiments, $\mathcal{H}_{\text{ss}}$ *physically* represents the state of the detector and $\mathcal{U}$ represents the time evolution of the system and detector during the measurement process. Since the state of the apparatus encodes the observed outcome, it is natural that the postmeasurement state ρ_ℓ given outcome ℓ recovers from applying $|\ell\rangle \langle \ell|_{\text{ss}}$ (see also Sec. 5). As a reminder, the derivations above apply to the *minimal* Stinespring representation with $\dim(\mathcal{H}_{\text{ss}}) = \mathcal{N}$. However, in general, there exist *nonminimal* Stinespring representations for which $\dim(\mathcal{H}_{\text{ss}}) > \mathcal{N}$. These generally correspond to the actual detector degrees of freedom in the experiment, and reduce to the minimal representation upon “binning” distinct states corresponding to the same outcome.

2.3.3. Single-qubit observables

First, consider measuring Z on a qubit. While different physical realizations of qubits may have different interpretations and require different implementations (see Sec. 4), they all share the

same minimal Stinespring representation (2.19). Since Z has eigenvalues ± 1 and eigenprojectors $P_{\pm} = (\mathbb{1} \pm Z)/2$, the measurement unitary is straightforward to work out,

$$\mathcal{U}_{[Z]} = \sum_{n=0,1} \frac{1}{2} (\mathbb{1} + (-1)^n Z) \otimes X_{\text{ss}}^n \quad (2.23a)$$

$$= |0\rangle\langle 0| \otimes \mathbb{1}_{\text{ss}} + |1\rangle\langle 1| \otimes X_{\text{ss}} = \text{CNOT}(\text{ph} \rightarrow \text{ss}), \quad (2.23b)$$

where $Z|n\rangle = (-1)^n |n\rangle$, so that the measurement unitary (2.23) is simply a CNOT gate with the physical system the “control” qubit and the detector the “target” qubit. As we discuss in Sec. 4, the detectors used in experiments are almost always more complicated than a single qubit; however, in the idealized limit of the measurement (i.e., vanishing probability of readout error), binning detector states into $n = 0$ versus $n = 1$ results in the minimal representation above (2.23).

We next generalize $\mathcal{U}$ (2.23) to arbitrary single-qubit observables $\mathcal{O}_j$ with eigenvalues $O_0 > O_1$; and decomposing this operator onto the Pauli basis for qubit j , we have that

$$\mathcal{O}_j \equiv \frac{1}{2} \text{tr}(\mathcal{O}_j) \mathbb{1}_j + \frac{1}{2} \sum_{\nu=x,y,z} \text{tr}(\sigma_j^\nu \mathcal{O}_j) \sigma_j^\nu, \quad (2.24)$$

so that there exists a traceless, involutory operator $\bar{\mathcal{O}}_j$ with the same eigenvectors [23] given by

$$\bar{\mathcal{O}}_j \equiv \frac{1}{O_0 - O_1} \sum_{\nu=x,y,z} \text{tr}(\mathcal{O}_j \sigma_j^\nu) \sigma_j^\nu \quad \text{so that} \quad P_m = \frac{1}{2} (\mathbb{1}_j + (-1)^m \bar{\mathcal{O}}_j), \quad (2.25)$$

whose eigenvalues are $\bar{O}_m = (-1)^m$, where P_m are the spectral projectors (2.1) for $\mathcal{O}_j$ (2.24) as well. The dilated unitary (2.19) that captures measurement of either $\mathcal{O}_j$ or its involutory part $\bar{\mathcal{O}}_j$ is

$$\mathcal{U}_{[\mathcal{O}_j]} = \frac{1}{2} (\mathbb{1}_j + \bar{\mathcal{O}}_j) \otimes \mathbb{1}_{\text{ss}} + \frac{1}{2} (\mathbb{1}_j - \bar{\mathcal{O}}_j) \otimes X_{\text{ss}}, \quad (2.26)$$

which applies to *any* single-qubit observable $\mathcal{O}_j$ (2.3) and any operator unitarily connected thereto.

2.3.4. Infinite spectra

Now, consider an observable $\mathcal{O}$ (2.3) with infinitely many eigenvalues ($\mathcal{N} \rightarrow \infty$). The Weyl operator X (2.20) is no longer defined, since “modulo $\mathcal{N}$ ” has no meaning. Although an alternative unitary to X (2.20) may be defined by labeling eigenvalues via $n \in \mathbb{Z}$ (instead of $n \in \mathbb{N}$) to realize a *minimal* Stinespring representation [31], we instead consider a *nonminimal* representation on

$$\mathcal{H}_{\text{ss}} = \bigotimes_{j \in \mathbb{N}} \mathbb{C}^2 \quad \text{with} \quad N_{\text{ss}} = \dim(\mathcal{H}_{\text{ss}}) = \infty, \quad (2.27)$$

corresponding to infinitely many qubits labeled by a “site” $j \in \mathbb{N}$. This nonminimal Stinespring Hilbert space (2.27) is not separable—its dimension is *uncountably* infinite by Cantor’s theorem.

Importantly, we note that the spectrum of $\mathcal{O}$ is *countably* infinite, so that the *minimal* Stinespring Hilbert space has a countable basis (i.e., is separable). It is common practice to define a separable analogue of $\mathcal{H}_{\text{ss}}$ (2.27) in terms of finite numbers of spin flips above the reference state $|\mathbf{0}\rangle_{\text{ss}}$ (with all qubits in the $+1$ eigenstate $|0\rangle$ of Z). However, measuring the boson occupation number may require exciting an infinite number of qubits; moreover, separability itself is not a concern. Instead, we note $\mathcal{H}_{\text{ss}}$ (2.27) can be partitioned into a countably infinite number of subspaces of configurations in which exactly n qubits are in the excited state $|1\rangle$, with all others in the default state $|0\rangle$. Each

subspace reflects a different outcome, and imposing the condition on $\mathcal{H}_{\text{ss}}$ (2.27) that all qubits in the state $|1\rangle$ are to the “left” of all qubits in the state $|0\rangle$ ensures that there is only one state from each subspace n —and thus, one state per outcome. In other words, labeling Stinespring qubits by the order in which they are excited leads to a countably, minimal Stinespring representation.

That minimal Stinespring Hilbert space is simply the bosonic Hilbert space $\ell^2(\mathbb{N}) \cong \mathbb{C}^\infty$. Still, $\mathcal{H}_{\text{ss}}$ (2.27) reflects the actual detector Hilbert space in experimental measurements of observables $\mathcal{O}$ with countably infinite spectra, as we discuss in Sec. 3. Initializing the detector qubits in the state $|i\rangle_{\text{ss}} = |\mathbf{0}\rangle_{\text{ss}}$, we arrive at the following extension of $\mathcal{U}$ (2.19) to countably infinite spectra,

$$\mathcal{U}_{[\mathcal{O}]} \equiv \sum_{n \in \mathbb{N}} P_n \otimes \prod_{j=1}^n X_{\text{ss},j}, \quad (2.28)$$

where $X_{\text{ss},j}$ is a Pauli X operator acting on the j th Stinespring qubit. Physically, this describes, e.g., the counting of the number of photons n at a given frequency by creating n excitations in a detector. Because the order of excitement is unimportant, all states with n excitations reflect the outcome O_m . We note that $\mathcal{U}$ (2.28) is only unitary on the full $\mathcal{H}_{\text{ss}}$ (2.27), and not on $\ell^2(\mathbb{N}) \subset \mathcal{H}_{\text{ss}}$.

2.3.5. Destructive measurements

We now consider *destructive* measurements, which are similar to the projective measurements considered thus far. While a projective measurement of $\mathcal{O}$ (2.3) in the state ρ resulting in the outcome O_m *projects* the system into the eigenstate $\rho_m \propto P_m \rho P_m$ of $\mathcal{O}$ (with the observed eigenvalue O_m), a destructive measurement *destroys* the measured property of the initial state ρ , generally resulting in a single postmeasurement state ρ' regardless of the observed eigenvalue O_m .

Consider a bosonic state $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ (e.g., a harmonic oscillator or electromagnetic mode), where a *projective* measurement of the boson number $N = a^\dagger a = \sum_{n=1}^{\infty} n |n\rangle\langle n|$ is captured by

$$\mathcal{U}_{[N]} : |\Psi_i\rangle = \sum_{n=0}^{\infty} c_n |n\rangle_{\text{ph}} \otimes |0\rangle_{\text{ss}} \mapsto |\Psi_f\rangle = \sum_{n=0}^{\infty} c_n |n\rangle_{\text{ph}} \otimes |n\rangle_{\text{ss}}, \quad (2.29)$$

under $\mathcal{U}_{[N]}$ (2.28). Following the measurement, if the Stinespring register is found to be in the state $|n\rangle_{\text{ss}}$ with n excited qubits, the system is in the state $|n\rangle_{\text{ph}}$ corresponding to exactly n bosons.

However, one could imagine that counting the number of bosons N requires destroying them one by one, until none remain. Such a *destructive* measurement is captured by the unitary

$$\mathcal{V}_{[N]} : |\Psi_i\rangle = \sum_{n=0}^{\infty} c_n |n\rangle_{\text{ph}} \otimes |0\rangle_{\text{ss}} \mapsto |\Psi'_f\rangle = \sum_{n=0}^{\infty} c_n |0\rangle_{\text{ph}} \otimes |n\rangle_{\text{ss}}, \quad (2.30)$$

where $|n\rangle_{\text{ss}}$ is a shorthand for *any* configuration of $\mathcal{H}_{\text{ss}}$ (2.27) with exactly n qubits in the state $|1\rangle$; the physical state following the destructive measurement is always empty.

Noting that $V : |0\rangle_{\text{ph}} \otimes |n\rangle_{\text{ss}} \mapsto |n\rangle_{\text{ph}} \otimes |0\rangle_{\text{ss}}$ is an isometry on $\mathcal{H}_{\text{ph}} \otimes \ell^2(\mathbb{N})$, we write

$$V^\dagger = \sum_{n=0}^{\infty} |0\rangle\langle n|_{\text{ph}} \otimes P_{\text{ss}}^{(n)} = \sum_{n=0}^{\infty} \left[a N^{-1/2} \right]^n \otimes \prod_{j=1}^n |1\rangle\langle 1|_j \prod_{k=n+1}^{\infty} |0\rangle\langle 0|_k, \quad (2.31)$$

where a is the *lowering operator* $a|n\rangle = \sqrt{n}|n-1\rangle$, $a^\dagger$ is the raising operator, and $N = a^\dagger a$. The isometry V realizes a *minimal* Stinespring representation on $\ell^2(\mathbb{N}) \subset \mathcal{H}_{\text{ss}}$ and satisfies

$$V^\dagger V = \mathbb{1}_{\text{ph}} \otimes \mathbb{1}_{\text{ss}} \quad \text{and} \quad V V^\dagger = \mathbb{1}_{\text{ph}} \otimes \mathbb{1}_{\text{ss}} - \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} |k\rangle\langle k|_{\text{ph}} \otimes P_{\text{ss}}^{(n)}, \quad (2.32)$$

as required. Moreover, $V^\dagger$ can always be embedded in a unitary $\mathcal{U}'$ acting on all of $\mathcal{H}_{\text{ss}}$ (see Sec. 3.2), so the destructive measurement of N is captured by the unitary operation

$$\mathcal{V}_{[N]} = \mathcal{U}' \mathcal{U}_{[N]}, \quad (2.33)$$

where $\mathcal{U}_{[N]}$ (2.28) realizes a projective measurement of N and $\mathcal{U}'$ is a unitary embedding of $V^\dagger$ (2.31) in $\mathcal{H}_{\text{dil}}$ (2.27). In this sense, projective and destructive measurements are equivalent up to some unitary $\mathcal{U}'$. In the case above, the destructive measurement is equivalent to a projective measurement followed by an outcome-dependent unitary $\mathcal{U}'$ that maps the postmeasurement physical state to some state $|0\rangle$; a projective measurement is equivalent to a destructive measurement followed by outcome-dependent restoration of the corresponding eigenstate $|n\rangle$.

2.4. von Neumann representation

An alternative to the unitary representation of measurements outlined in Sec. 2.3 is von Neumann's formulation in terms of a Hamiltonian H [32, 33]. Like the unitary $\mathcal{U}$ (2.17), the Hamiltonian H acts on the *dilated* Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H}_{\text{ph}} \otimes \mathcal{H}_{\text{ss}}$, where now, $\mathcal{H}_{\text{ss}}$ explicitly represents the state of the measurement apparatus. The Hamiltonian formulation is often more useful in the context of quantum optics [37, 38]. We review the von Neumann representation in Sec. 2.4.1, consider the example of the Stern-Gerlach experiment [39–41] in Sec. 2.4.2, and connect the pointer Hamiltonian H to the Stinespring unitary $\mathcal{U}$ (2.17) in Sec. 2.4.3.

2.4.1. Pointer Hamiltonian

We now review von Neumann's description of measurements in terms of a “pointer particle” [32, 33]. This model represents an early attempt to account for the details of the measurement apparatus, albeit in terms of H and not the unitary $\mathcal{U}$ (2.19) it generates. The pointer particle propagates freely, apart from a coupling between the observable $\mathcal{O}$ to be measured and the momentum p of the pointer particle (with coupling strength λ), i.e.,

$$H = H_0 \otimes \mathbb{1}_{\text{pp}} + \mathbb{1}_{\text{ph}} \otimes \frac{p^2}{2M} + \lambda \mathcal{O} \otimes p, \quad (2.34)$$

where H_0 is the Hamiltonian for the physical system alone, and $M \gg 1$ is the mass of the pointer particle (“pp”). As a comment, the Wigner-Araki-Yanase Theorem [42, 43] generally requires that $\mathcal{O}$ commute with H_0 or else the uncertainty is generally guaranteed to be large.

The combined system is prepared in the state $|\Phi(0)\rangle = |\phi(0)\rangle_{\text{ph}} \otimes |\varphi(0)\rangle_{\text{pp}}$, with

$$|\phi(0)\rangle_{\text{ph}} = \sum_{m=0}^{\mathcal{N}-1} \sum_{\ell=1}^{r_m} \phi_{m,\ell} |m, \ell\rangle_{\text{ph}} \quad (2.35a)$$

$$|\varphi(0)\rangle_{\text{pp}} = \int_{\mathbb{R}} dx \varphi_0(x) |x\rangle_{\text{pp}} = \int_{\mathbb{R}} dx \frac{\exp(-(x-x_0)^2/4\sigma^2)}{(2\pi\sigma^2)^{1/4}} |x\rangle_{\text{pp}}, \quad (2.35b)$$

where $\mathcal{O} |m, \ell\rangle = O_m |m, \ell\rangle$ and ℓ runs over the r_m degenerate states with eigenvalue O_m . The state $|\Phi(0)\rangle = |\phi(0)\rangle_{\text{ph}} \otimes |\varphi(0)\rangle_{\text{pp}}$ satisfies $\langle x \rangle = x_0$, $\langle p \rangle = 0$, $\Delta x = \sigma$, and $\Delta p = \hbar/2\sigma$.

Assuming that $[\mathcal{O}, H_0] = 0$ and $M \gg 1$, one can ignore both H_0 and the pointer's dispersion in Eq. 2.34. In this limit, the time evolution operator is well approximated by the unitary

$$\mathcal{U}_{\text{eff}}(t) = \sum_{m=0}^{\mathcal{N}-1} P_m \otimes e^{-it\lambda O_m p/\hbar}, \quad (2.36)$$

under which, the initial state $|\Phi(0)\rangle$ (2.35) evolves into

$$\begin{aligned}
|\Phi(t)\rangle &= \mathcal{U}_{\text{eff}}(t) |\Phi(0)\rangle = \mathcal{U}_{\text{eff}}(t) \sum_{m=0}^{\mathcal{N}-1} \sum_{\ell=1}^{n_m} \int_{\mathbb{R}} dx \phi_{m,\ell} |m, \ell\rangle_{\text{ph}} \otimes \varphi_0(x) |x\rangle_{\text{pp}} \\
&= \sum_{m,k=0}^{\mathcal{N}-1} \sum_{\ell=1}^{n_m} \int_{\mathbb{R}} dx (\phi_{m,\ell} P_k |m, \ell\rangle_{\text{ph}}) \otimes \left(e^{-\lambda t O_k} \partial_x \varphi_0(x) \right) |x\rangle_{\text{pp}} \\
&= \sum_{m=0}^{\mathcal{N}-1} \sum_{\ell=1}^{n_m} \int_{\mathbb{R}} dx \phi_{m,\ell} |m, \ell\rangle_{\text{ph}} \otimes \varphi_0(x - \lambda t O_m) |x\rangle_{\text{pp}}, \tag{2.37}
\end{aligned}$$

which can be summarized as the following map (2.36) on the initial state (2.35),

$$|\phi(0)\rangle_{\text{ph}} \otimes \int_{\mathbb{R}} dx \varphi_0(x) |x\rangle_{\text{pp}} \mapsto \sum_{m=0}^{\mathcal{N}-1} (P_m |\phi(0)\rangle)_{\text{ph}} \otimes \int_{\mathbb{R}} dx \varphi_0(x - \lambda t O_m) |x\rangle_{\text{pp}}, \tag{2.38}$$

meaning that, for $t > 0$, the full state is a sum over outcomes m of the product of the physical postmeasurement state $P_m |\phi(0)\rangle_{\text{ph}}$ and the pointer state $\varphi_0(x - \lambda t O_m)$.

In other words, the position of the pointer particle indicates the measurement outcome: If the initial uncertainty $\Delta x = \sigma$ in the pointer's position x satisfies $\sigma \ll \lambda t \min(\delta O_m)$, where $\min(\delta O_m)$ is the minimum difference between consecutive eigenvalues of $\mathcal{O}$ (ordered by absolute value), then one can distinguish the particular eigenvalue O_m by the displacement of the pointer particle's wave packet initially centered about x_0 with standard deviation $\Delta x = \sigma$ (2.35b).

The final state (2.38) is realized in the limit $[H_0, \mathcal{O}] = 0$ and $M \rightarrow \infty$ in Eq. 2.34. When $[H_0, \mathcal{O}] \neq 0$ or the pointer particle's mass is small, the same result (2.38) can be achieved by increasing the coupling to the pointer particle ($\lambda \gg 1$) and decreasing the interaction time ($t \ll 1$), so that H_0 and the dispersion term $p^2/2M$ in Eq. 2.34 do not have time to introduce noise to Eq. 2.38. Finally, we comment that for any fixed t , the von Neumann unitary evolution operator $\mathcal{U}(t)$ (2.36) is equivalent to the Stinespring measurement unitary $\mathcal{U}$ (2.17).

2.4.2. The Stern-Gerlach experiment

One of the first examples of a quantum measurement was realized by Stern and Gerlach [39–41, 44]. A stream of particles with intrinsic magnetic moments (e.g., silver atoms) travels in the positive x direction through an apparatus that applies a magnetic field with nonzero gradient in the z direction. The particles have an internal spin magnetic moment $\boldsymbol{\mu}$, resulting in a potential $V = -\boldsymbol{\mu} \cdot \mathbf{B}$; since the gradient is nonvanishing in the z direction, we expect a nonzero force on the particle, proportional to $\mu = |\boldsymbol{\mu}|$. In this sense, the atom is *its own pointer particle*: The position of the atom after exiting the apparatus effectively measures its magnetic moment $\boldsymbol{\mu}$ in the z direction. Crucially, while classical mechanics predicts a continuum of final positions of the particles, corresponding to a continuum of possible magnetic moments $\mu_z = |\mu| \cos(\theta)$, the actual experiment shows *quantized* outcomes, corresponding to the particle appearing at $z = \pm \delta z$. The value of δz is determined by experimental details that do not vary between shots.

We first construct the Hamiltonian describing the particles inside the apparatus (i.e., from $x = 0$ to $x = L$). Assuming the magnetic field is $\mathbf{B} = (0, -by, B_0 + bz)$, where $B_0 \gg b$, we have that

$$H_{\text{SG}} = -\frac{\hbar^2}{2M} (\partial_x^2 + \partial_y^2 + \partial_z^2) + \mu_B (-by \sigma^y + (B_0 + bz) \sigma^z), \tag{2.39}$$

where σ^n is the n th Pauli matrix, and the particles' magnetic moment has vector components $\mu_n = \mu_B g \sigma^n / 2 \approx \mu_B \sigma^n$, where μ_B is the Bohr magneton. Instead of the time-dependent Schrödinger equation under (2.39), we consider the more straightforward Heisenberg evolution of operators.

We make the standard assumption that the initial state is separable with respect to the coordinates x, y, z . Noting that x is decoupled, we assume that $\phi_x(p_x) \propto \exp(-(p_x - Mv)^2 / 4\delta_x^2)$, where δ_x is the x -momentum variance, and v is a velocity. In the Heisenberg picture, we find

$$x(t) = x(0) + \frac{t}{M} p_x(0) \implies \langle x \rangle_t = \langle \phi_x(p_x) | x(t) | \phi_x(p_x) \rangle = vt, \quad (2.40)$$

in accordance with Ehrenfest's Theorem [45]. Hence, the particle traverses the apparatus in time $t = L/v$; we assume that the particle's z position is measured upon exiting the apparatus at $x = L$.

The analogous expressions to Eq. 2.40 for y and z are far more complicated, and do not truncate at finite order in t . Expressions up to $O(t^5)$ appear in App. A.1, and suggest simplifications. First, since the x -dependent part of $|\Psi(t)\rangle$ evolves independently, it may be considered separately. Second, we find in App. A.1 that $\langle y \rangle_t = \langle \Psi(0) | y(t) | \Psi(0) \rangle$ is identically zero to $O(t^5)$ in generic initial states, and we expect this holds for most or all orders, so that the dynamics in the y direction can be ignored (i.e., we can fix $y = 0$). As a result, H_{SG} (2.39) is well approximated by

$$H_z = -\frac{\hbar^2}{2M} \partial_z^2 + \mu_B (B_0 + bz) \sigma^z, \quad (2.41)$$

where we ignore y dynamics. Suppose that the particle is initialized in the state

$$|\Psi(0)\rangle = \sum_{s=\pm 1} \int_{\mathbb{R}} dz \Psi_s(z, 0) |z, s\rangle = \sum_{s=\pm 1} c_s \int_{\mathbb{R}} dz \frac{e^{-(z-z_0)^2/4\delta^2}}{(2\pi\delta^2)^{1/4}} |z, s\rangle, \quad (2.42)$$

corresponding to a Gaussian wave packet centered at $z = z_0$ with variance δ .

The unitary operator that generates time evolution under H_z (2.41) is given by

$$\begin{aligned} \mathcal{U}(t) &= \exp\left(\frac{t\mu_B}{i\hbar} (B_0 + bz) \sigma^z + \frac{i\hbar t}{2M} \partial_z^2\right) \\ &= e^{ib^2\mu_B^2 t^3/M\hbar} \exp\left(\frac{t\mu_B}{i\hbar} (B_0 + bz) \sigma^z\right) \exp\left(\frac{i\hbar t}{2M} \partial_z^2\right) \exp\left(-\frac{b\mu_B}{M} t^2 \partial_z \sigma^z\right), \end{aligned} \quad (2.43)$$

via the Zassenhaus formula [46]. Applying this operator to the initial state $|\Psi(0)\rangle$ (2.42) leads to

$$\Psi_s(z, t) = c_s \frac{e^{ib^2\mu_B^2 t^3/M\hbar} e^{-ist\mu_B(B_0+bz)/\hbar}}{(2\pi)^{1/4} (\delta + \frac{i\delta t}{2\delta})^{1/2}} \exp\left[-\frac{\left(z - z_0 - \frac{b\mu_B}{M} s t^2\right)^2}{4\delta^2} \frac{1}{1 + \frac{i\hbar t}{2M\delta^2}}\right], \quad (2.44)$$

as derived in detail in App. A.2. The probability distribution for the particle's z coordinate is

$$p_s(z, t) = p(s) \mathcal{N}\left(z_0 + \frac{b\mu_B}{M} s t^2, \delta^2 \left(1 - \frac{\hbar^2 t^2}{4M^2 \delta^4}\right)\right), \quad (2.45)$$

where $p(s)$ is the probability of $\langle \sigma^z \rangle = s = \pm 1$ in the initial state (2.42), and $\mathcal{N}(z_{\text{av}}, \sigma^2)$ is a normal distribution with mean z_{av} and variance σ^2 . We evaluate $p_s(z, t)$ (2.45) at time $t = L/v$, where v is the initial x velocity and L is the length of the apparatus in the x direction.

Evolving the particle's z coordinate under H_z (2.41) in the Heisenberg picture leads to

$$z(t) = z + \frac{t}{M} p_z - \frac{b\mu_B}{2M} t^2 \sigma^z \implies \langle z \rangle_{s,t} = \langle \Psi_s(0) | z(t) | \Psi_s(0) \rangle = z_0 - \frac{b\mu_B}{2M} s t^2, \quad (2.46)$$

evaluated in $|\Psi(0)\rangle$ (2.42), where the x dynamics (2.40) determine $t = L/v$.

Because the particle’s position is classically accessible, one expects to observe the classical behavior associated with the expectation values of x (2.40) and z (2.46), with y trivial. However, the particle’s magnetic moment (captured via s) is not classically observable; moreover, one cannot observe a superposition of the macroscopically distinct z -basis states—corresponding to the two spin values $s = \pm 1$ —in experiment (by nature, classical objects cannot appear in macroscopically distinct superpositions). Instead, any experiment appears to result in a shift of z by *either* $-\alpha$ (when $s = 1$) or $+\alpha$ (when $s = -1$). Importantly, there is not a continuum of favored displacements, but a *single* displacement, with sign given by the particle’s internal magnetic moment with eigenvalues $s = \pm 1$.

2.4.3. Connection to the Stinespring unitary

We now connect von Neumann’s Hamiltonian representation to the Stinespring unitary $\mathcal{U}$ (2.17) from Sec. 2.3. Both representations act on the *dilated* Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H}_{\text{ph}} \otimes \mathcal{H}_{\text{ss}}$ (2.8). In the unitary case, $\mathcal{H}_{\text{ss}}$ is a bookkeeping tool that stores the measurement outcome; in the Hamiltonian case, it corresponds physically to an observable property of the “pointer” particle [32, 33].

The pointer Hamiltonian H (2.34) generates the unitary evolution operator $\mathcal{U}(t) = \exp(-itH/\hbar)$ (2.36). Applying $\mathcal{U}(t)$ entangles the system and pointer particle, resulting in a sum over measurement outcomes m of the physical state $P_m|\psi\rangle$ and the pointer state $|x_0 - \alpha O_m\rangle$. In this sense, the von Neumann unitary $\mathcal{U}(t)$ (2.36) *is* the Stinespring unitary $\mathcal{U}$ (2.17), though possibly nonminimal, as in the case of the Stern-Gerlach experiment [39–41] discussed in Sec. 2.4.2. However, recording a particular outcome requires “binning” the states of the pointer particle into a minimal set of $\mathcal{N}$ “Stinespring states.” In the Stern-Gerlach experiment, taking $z_0 = 0$ and for fixed initial x velocity v and magnetic field $B_z = B_0 + bz$, there are possible final z positions of the pointer, localized to $z_f = \pm b\mu_B L^2/Mv^2$, up to experimental imprecision and variations encoded in the initial state (2.42). These are binned easily binned into $s = \pm$ according to $s = \text{sgn}(z_f)$.

Thus, we expect that the Stinespring unitary $\mathcal{U}$ (2.17) is *not* merely a bookkeeping tool, but a *physical* operator corresponding to the time evolution of the system and measurement apparatus (a detector, pointer particle, etc.) during the measurement process. The Stinespring Hilbert space $\mathcal{H}_{\text{ss}}$ gives a (minimal) representation of the state of the apparatus, with binning of states already encoded. We stress that the binning is itself part of the measurement; we explore numerous examples in which the $\dim(\mathcal{H}_{\text{ss}}) \gg \mathcal{N}$ in Secs. 3 and 4. In fact, this is to be expected when $\mathcal{N} \rightarrow \infty$ (see Sec. 2.3.4), and the minimal Stinespring representation implicitly encodes equivalence classes between states of $\mathcal{H}_{\text{ss}}$ that reflect the same measurement outcome. The technical details of how outcomes are extracted from the state of the detector are generally of little importance from the perspectives of (i) a theory of projective and destructive measurements and (ii) the practical or analytical treatment of protocols involving measurements and outcome-dependent operations [21, 22].

2.5. Summary of measurement representations

Before moving on to experimental implementations, we briefly review the mathematical representations of projective and destructive measurements developed thus far. The axioms of quantum mechanics [1–8] dictate that such a measurement of an observable $\mathcal{O}$ (2.3) in a system described by a density matrix ρ results in one of the $\mathcal{N}$ unique eigenvalues O_m of $\mathcal{O}$, such that

$$\Pr(O_m) \equiv p_m = \text{tr}(P_m \rho) \quad \text{and} \quad \rho \mapsto \Phi_m(\rho) = p_m \rho_m = P_m \rho P_m, \quad (2.47)$$

where ρ_m is the postmeasurement density matrix given outcome m and the eigenprojector P_m (2.1) satisfies $\mathcal{O}P_m = O_m P_m$ and Eq. 2.2. Averaging over outcomes leads to

$$\rho \mapsto \Phi(\rho) = \rho_{\text{av}} \equiv \sum_{m=0}^{\mathcal{N}-1} p_m \rho_m = \sum_{m=0}^{\mathcal{N}-1} P_m \rho P_m, \quad (2.48)$$

which is a quantum channel (a CPTP map) [24], with $\Phi(\rho) = \sum \Phi_m(\rho)$.

The *Kraus representations* (2.11) of the quantum operation $\Phi_m : \rho \mapsto p_m \rho_m$ (2.47) and the related quantum channel $\Phi : \rho \mapsto \rho_{\text{av}}$ (2.48) are given straightforwardly by [20]

$$\Phi_m(\rho) = E_m \rho E_m^\dagger, \quad \Phi(\rho) = \sum_{m=0}^{\mathcal{N}-1} E_m \rho E_m^\dagger, \quad \text{where } E_m = P_m, \quad (2.49)$$

so that the *Kraus operator* E_m is simply the projector P_m (2.1), meaning that

$$\sum_{m=0}^{\mathcal{N}-1} E_m^\dagger E_m = \sum_{m=0}^{\mathcal{N}-1} P_m = \mathbb{1}, \quad (2.50)$$

i.e., the Kraus operators form a complete set (2.13). Quantum channels are trace preserving, and involve a complete set of Kraus operators, while quantum operations are trace decreasing and involve a proper subset of the complete set of Kraus operators that define an associated channel.

Related to the Kraus representation (2.49) of quantum operations and channels is a *unitary representation* (2.17), first proven by Kraus using the Stinespring dilation theorem for finite-dimensional systems [17–20, 28, 30]. An extension to infinite-dimensional systems is the subject of forthcoming work [31]. The unitary “Stinespring” representation is captured by

$$\Phi(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger \right) \quad \text{and} \quad \Phi_m(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger P_m^{\text{ss}} \right), \quad (2.51)$$

where $|i\rangle_{\text{ss}}$ is some “default” initial state on $\mathcal{H}_{\text{ss}} = \mathbb{C}^{\mathcal{N}}$, P_m^{ss} projects onto $|m\rangle_{\text{ss}} \in \mathcal{H}_{\text{ss}}$, and $\mathcal{U}$ acts on $\mathcal{H}_{\text{dil}} = \mathcal{H} \otimes \mathcal{H}_{\text{ss}}$ (2.8) and can be expressed in terms of the Kraus operators (2.49) via

$$\mathcal{U} |\psi\rangle_{\text{ph}} \otimes |0\rangle_{\text{ss}} = \sum_{m=0}^{\mathcal{N}-1} (E_m |\psi\rangle)_{\text{ph}} \otimes |m\rangle_{\text{ss}} \quad \implies \quad \mathcal{U} = \sum_{m=0}^{\mathcal{N}-1} P_m \otimes X_{\text{ss}}^m, \quad (2.52)$$

for projective measurements, where $X^m|0\rangle = |m\rangle$ is the $\mathcal{N}$ -state Weyl shift operator (2.20), and we have taken $|i\rangle_{\text{ss}} = |0\rangle_{\text{ss}}$ without loss of generality. Meanwhile, *destructive measurements* are represented by a dilated unitary $\mathcal{V}$ (2.33) that is equal to the projective-measurement unitary $\mathcal{U}$ (2.52) by another dilated unitary $\mathcal{U}'$. For either type of measurement, $\mathcal{U}$ (2.52) realizes a *minimal* Stinespring representation [27], where $\dim(\mathcal{H}_{\text{ss}}) = \mathcal{N}$ is the number of unique eigenvalues of the measured observable $\mathcal{O}$ (2.3). However, there also exist nonminimal representations with $\dim(\mathcal{H}_{\text{ss}}) > \mathcal{N}$, which reduce to Eq. 2.52 upon “binning” different states of $\mathcal{H}_{\text{ss}}$ corresponding to the same outcome O_m . We discuss how to extract various statistics from $\mathcal{U}$ (2.52) in Sec. 5.

Finally, the *von Neumann representation* describes measurements using a Hamiltonian H_{vN} (2.34), which acts on a dilated Hilbert space $\mathcal{H}_{\text{dil}}$ (2.8) that includes the physical system and a “pointer” particle [32]. The pointer plays the role of the detector, and H_{vN} relates to $\mathcal{U}$ (2.51) via

$$\mathcal{U} = \exp \left(-\frac{i}{\hbar} \int_0^t d\tau H_{\text{vN}}(\tau) \right), \quad (2.53)$$

i.e., $\mathcal{U}$ (2.51) corresponds to evolution under the von Neumann Hamiltonian $H_{\text{vN}}(\tau)$ (2.34) for time t , and connects the Kraus and von Neumann representations of measurements. In general, H_{vN} (2.34) leads to a *nonminimal* Stinespring representation; the minimal representation (2.51) recovers upon identifying subspaces of $\mathcal{H}_{\text{ss}}$ that reflect the same outcome. In the Stern-Gerlach experiment [39–41, 44], e.g., the minimal representation has $\mathcal{H}_{\text{ss}} = \mathbb{C}^2$, corresponding to states with states where $s = \text{sgn}[z(t) - z(0)] = \pm 1$ for any values of M , L , v , and b . The distinct subspaces with $s = \pm 1$ reflect the (quantized) z component of the atom’s intrinsic spin.

We discuss particular realizations of $\mathcal{U}$ and/or H_{vN} for measurements of photons and qubits in Secs. 3 and 4, respectively. While a dilated unitary representation $\mathcal{U}$ (2.51) of measurements is guaranteed for any axiomatic formulation of quantum mechanics [1–8], this only implies that $\mathcal{U}$ is a valid bookkeeping tool. However, the connection (2.53) to von Neumann’s pointer Hamiltonian suggests that $\mathcal{U}$ is, in fact, *physical*. In Secs. 3 and 4, we show how this is the case in several of the most common implementations of projective and destructive measurements, and posit that this holds in all cases. In Sec. 5, we discuss how to use the Stinespring representation to describe generic quantum protocols, as well as implications of the unitary $\mathcal{U}$ (2.51) being physical.

3. PHOTON MEASUREMENTS

We first consider the measurement of photons, which not only diagnose optical systems, but are relevant to the measurements of atomic and molecular systems that we discuss in Sec. 4. As noted in Sec. 2.3.5, one generally expects the counting of photons to be *destructive*, rather than projective—though we comment that there exist projective measurements of photon number [47–49] naturally described by Eq. 2.28. Importantly, such destructive measurements are nonetheless captured by the von Neumann (2.34) and Stinespring (2.33) representations. These representations require that we model the measurement apparatus explicitly, in contrast to the device-independent description of photodetection introduced by Glauber [50].

This section is organized as follows. In Sec. 3.1, we review the description of photons as oscillator-like excitations of the electromagnetic fields in matter-free regions, and then extend to optical modes in cavities. In Sec. 3.2, we consider the destructive measurement of photon number, which serves as a proxy for measurements of intensity and other properties related to energy or occupation number [50–52]. In Sec. 3.3 we consider homodyne detection [53–57] of photon quadratures (3.14), which relates to general interferometric probes [37, 38, 53–56].

3.1. Electromagnetic modes

Quantum mechanically, photons are the excitations of the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$, which in a matter-free region are described by the continuum Hamiltonian [44],

$$H_{\text{EM}} = \frac{\epsilon_0}{2} \int_V d^3\mathbf{x} \left(\mathbf{E}^2(\mathbf{x}) + c^2 \mathbf{B}^2(\mathbf{x}) \right), \quad (3.1)$$

where $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\partial_t \mathbf{A}$, and we use the traverse (Coulomb) gauge $\nabla \cdot \mathbf{A} = 0$.

Regarding H_{EM} (3.1), we interpret the electric field $\mathbf{E} = -\partial_t \mathbf{A}$ as the “momentum” conjugate to $\mathbf{A}$, and the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ as a “gradient” of $\mathbf{A}$. Then, H_{EM} (3.1) resembles a theory of bosonic excitations (e.g., in a harmonic solid) and, in a box with volume V , we write

$$A_j(\mathbf{x}) \equiv \sqrt{\frac{\hbar}{2c\epsilon_0 V}} \sum_{\mathbf{k}} |\mathbf{k}|^{-1/2} \sum_{s=1,2} \left(e^{i\mathbf{k}\cdot\mathbf{x}} \hat{e}_{s,j}(\mathbf{k}) a_{\mathbf{k},s} + e^{-i\mathbf{k}\cdot\mathbf{x}} \hat{e}_{s,j}^*(\mathbf{k}) a_{\mathbf{k},s}^\dagger \right), \quad (3.2)$$

where $\mathbf{k}$ is a Fourier wavevector, the ladder operators obey the canonical commutation relation

$$\left[a_{\mathbf{k}',s'}, a_{\mathbf{k},s}^\dagger \right] = \delta_{s,s'} \delta(\mathbf{k} - \mathbf{k}') , \quad (3.3)$$

and $\hat{e}_{s,i}(\mathbf{k})$ is the i th component of the *polarization* unit vector $\hat{e}_s$, which satisfies

$$\sum_{s=1,2} \hat{e}_{s,i}(\mathbf{k}) \hat{e}_{s,j}^*(\mathbf{k}) = \delta_{i,j} - \frac{k_i k_j}{|\mathbf{k}|^2} , \quad (3.4)$$

so that the polarization vectors span the plane perpendicular to $\mathbf{k}$, and there are two modes $s = 1, 2$ instead of the three associated with the photon's spin 1 due to gauge redundancy.

Using Eq. 3.2 for $\mathbf{A}$ (and thereby $\mathbf{E}$), we express H_{EM} in terms of ladder operators (3.3),

$$H_{\text{EM}} = \sum_{\mathbf{k}} \sum_{s=1,2} \hbar \omega(\mathbf{k}) \left(a_{\mathbf{k},s}^\dagger a_{\mathbf{k},s} + \frac{1}{2} \right) , \quad (3.5)$$

where $N_{\mathbf{k},s} = a_{\mathbf{k},s}^\dagger a_{\mathbf{k},s} \in \mathbb{N}$ is the number of photons with polarization s and wavevector $\mathbf{k}$, with corresponding dispersion $\omega(\mathbf{k}) = c|\mathbf{k}|$ [44]. These occupation numbers fully characterize the electromagnetic fields in a matter-free box of volume V .

In many scenarios, the spatial structure of the electromagnetic fields is modified by the presence of mirrors (as in the case of cavities) or some dense medium. In other cases, the range of relevant frequencies is sufficiently small that the factor of $k^{-1/2} \propto \omega^{-1/2}$ can be pulled out of the sum in Eq. 3.2 to good approximation [37, 38]. These cases are still described by H_{EM} (3.5), but the expansion of $\mathbf{A}$ in Eq. 3.2 is neither appropriate nor convenient. Instead, one writes

$$\mathbf{A}(\mathbf{x}) \equiv \sqrt{\frac{\hbar}{2\epsilon_0 c}} \sum_n \sum_{s=1,2} \left(\mathbf{u}_{n,s}(\mathbf{x}) a_{n,s} + \mathbf{u}_{n,s}^*(\mathbf{x}) a_{n,s}^\dagger \right) , \quad (3.6)$$

where the vector fields $\mathbf{u}_{n,s}(\mathbf{x})$ are (normalized) classical solutions to Maxwell's equations with the appropriate boundary conditions. Most of these optical “modes” propagate and resemble plane waves far from any mirrors or media, but some may be localized in space (e.g., within a cavity). Each mode n, s is associated with the creation and annihilation operators $a_{n,s}^\dagger$ and $a_{n,s}$, respectively, as with the plane-wave photon modes that realize in free space (3.3) [37, 38].

We also comment that, while a full accounting of all electromagnetic modes and the precise form of $\mathbf{A}$ (3.2) is needed to resolve the spatiotemporal profile of the electromagnetic fields, it is not necessary to the measurements of interest herein [37, 38]. Instead, it is generally sufficient to consider one or two electromagnetic modes, which may correspond to a particular wavevector and polarization in free space, a particular confined mode in a cavity, or similar.

3.2. Destructive measurement of photon number

As alluded to in Sec. 2.3.5, measuring the number of photons at wavevector $\mathbf{k}$ (3.5)—or, more generally, the occupation or intensity of an electromagnetic mode n (3.6)—is generally *destructive*, rather than projective. The reason is that the detector counts the photons by absorbing them, destroying the system's state. However, such measurements are nonetheless captured by a von Neumann Hamiltonian H (2.34) or Stinespring unitary $\mathcal{V}$ (2.30) acting on $\mathcal{H}_{\text{dil}}$ (2.8), and the latter is equivalent to a projective measurement $\mathcal{U}$ (2.28) followed by a dilated unitary $\mathcal{U}'$ (2.33).

We now consider a simplified description of the destructive measurement of occupation number (i.e., photon counting), based on a toy model of a photodetector comprising $\mathcal{N}$ two-level systems.

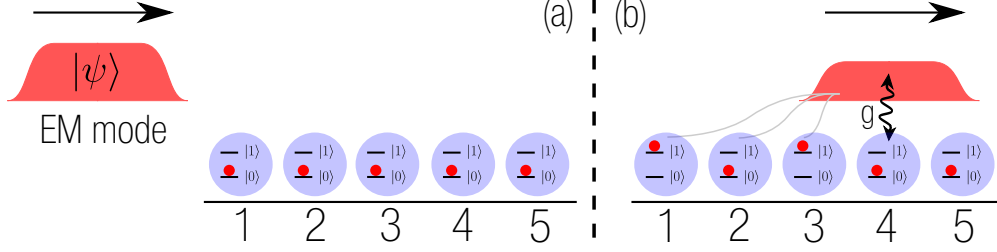


FIG. 1. A simple mode of destructive measurements of photon number or intensity. (a) Prior to the measurement, the electromagnetic mode realizes a mildly localized wave packet and the electrons in the detector are represented by $\mathcal{N}$ two-level electrons on a line, indexed k . (b) During the measurement, the photons interact with the electrons (strength g) in sequence; each absorption exciting an electron to the state $|1\rangle$, leading to a current proportional to the number of absorbed photons.

This toy model schematically captures a variety of distinct physical processes by which photons are detected in actual experiments. Destructive counting of photons necessarily requires their absorption, which transfers energy and momentum to the detector in the form of an “excitation” ($|0\rangle \mapsto |1\rangle$). The number of excitations reflects the number of photons counted; for the result to be readable, the excitations must result in a classically discernible signal. For example, the photons may excite electrons to the conduction band, producing a detectable current; electron-hole pairs in a semiconductor, altering the material’s conductivity; phonons in a lattice, altering the material’s thermal properties; and other mechanisms [58].

Importantly, all of these experimental realizations of photon counting are captured by time evolution on a dilated Hilbert space via a Hamiltonian (2.34) or unitary (2.17). In the interest of clarity, we neglect many technical details related to sources of noise, “dark counts” (i.e., false positives), thermal effects, differences between different detector realizations, the presence of multiple electromagnetic modes, and so on. Such considerations are discussed in great detail in the literature [58]. We also comment that, while some detectors of electromagnetic radiation are not well suited to counting individual photons, they are nonetheless suitable for coarse-grained measurements of intensity. The simplified model of photon counting we consider below is also a convenient model for intensity measurements [51] and similar energetic measurements. While the precise details of the detector degrees of freedom, set of outcomes, binning of states into outcomes, and other details may differ, the minimal Stinespring unitary is the same.

With these caveats in mind, we now consider the simplified model of destructive measurement of photon number [59, 60]. Suppose that the detector consists of electrons in the default state $|0\rangle$, which excite to the conducting state $|1\rangle$ upon absorption, giving rise to a classically discernible electrical current (though this model captures generic photodetectors). We also restrict to a *single* electromagnetic mode, with ladder operators a and $a^\dagger$. It is useful to picture this mode as a propagating wavepacket—i.e., a localized mode moving in space towards the detector [see Fig. 1(a)].

The $\mathcal{N}$ detector electrons are described by the fermion operators $f_{s,k}$, where $s = 0, 1$ labels the “orbital” and $k \in \{1, 2, \dots, \mathcal{N}\}$ reflects the order of excitation [see Fig. 1(a)]. The initial states of the electromagnetic mode and detector are, respectively, $|\psi\rangle \in \ell^2(\mathbb{N})$ and $|i\rangle_{ss} = |\mathbf{0}\rangle = |000 \cdots 000\rangle$. The light interacts with the k th electron for a time τ under the Hamiltonian

$$H_k = i g \left(a \otimes f_{1,k}^\dagger f_{0,k} - a^\dagger \otimes f_{0,k}^\dagger f_{1,k} \right), \quad (3.7)$$

where g is the strength of the electron-photon coupling. Ignoring electron-electron interactions, we replace the fermion operators with spin operators according to $f_{1,k}^\dagger f_{0,k} \equiv \sigma_k^+$ and $f_{0,k}^\dagger f_{1,k} \equiv \sigma_k^-$.

The sequential interaction between the optical mode and the effective spins 1/2 leads to

$$|\Psi\rangle = e^{g\tau(a\sigma_N^+ - a^\dagger\sigma_N^-)} \dots e^{g\tau(a\sigma_1^+ - a^\dagger\sigma_1^-)} |\psi\rangle \otimes |\mathbf{0}\rangle_{\text{ss}}, \quad (3.8)$$

where $|\Psi\rangle \in \mathcal{H}_{\text{dil}}$ is final state of the light and detector. We also define the parameter

$$\zeta \equiv \mathcal{N} g^2 \tau^2, \quad (3.9)$$

to be fixed and finite as we take the limits $g\tau \ll 1$ and $\mathcal{N} \gg 1$, so that Eq. 3.8 becomes

$$|\Psi\rangle = \exp\left[-\frac{\zeta}{2} a^\dagger a\right] \exp\left[(1 - e^{-\zeta})^{1/2} a \otimes B^\dagger\right] |\psi\rangle \otimes |\mathbf{0}\rangle_{\text{ss}}. \quad (3.10)$$

A careful derivation of the new ladder operator $B^\dagger$ appears in App. B and results in

$$B^\dagger \equiv \frac{1}{\sqrt{\mathcal{N}}} \frac{\sqrt{\zeta}}{\sqrt{e^\zeta - 1}} \sum_{k=1}^{\mathcal{N}} e^{\zeta(\mathcal{N}-k)/(2\mathcal{N})} \sigma_k^+, \quad (3.11)$$

as a collective raising operator on $\mathcal{H}_{\text{ss}}$ with $[B, B^\dagger] = 1$ in the limit $\mathcal{N} \rightarrow \infty$ and $\tau \rightarrow 0$, with leading corrections proportional to $g\tau$ and $\mathcal{N}^{-1/2}$ (see App. B for details).

The operator $B^\dagger$ (3.11) creates bosonic excitations in the detector above the “vacuum” state $|\mathbf{0}\rangle_{\text{ss}} = |000\dots 000\rangle$. Note that corrections to the final state $|\Psi\rangle$ (3.10) vanish as $\mathcal{O}(\mathcal{N}^{-1/2})$ as $\mathcal{N} \rightarrow \infty$. Hence, the electronic state of the detector is well approximated by a bosonic mode, where $B^\dagger B$ counts the number of electrons in the state $|1\rangle$ after interacting with the light. Those excited electrons produce a classically detectable current proportional to the number of absorbed photons. While other means of counting photons exist, they generically involve converting photons into excitations, leading to a classical signal from which the number of photons can be inferred.

The *attenuation parameter* ζ (3.9) controls the efficiency of the transduction of photons into excitations of the detector; its name refers to the “attenuation” of electromagnetic energy with each absorbed photon. The transduction—and thus, the counting—of photons is most efficient when the attenuation parameter $\zeta \gg 1$ is large but finite as $\mathcal{N} \rightarrow \infty$ and $g\tau \rightarrow 0$. For example, if the system is initially in the Fock state $|\psi_0\rangle_{\text{ph}} = |n\rangle$ with exactly n photons, then the measurement results in

$$|n\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} \mapsto |\Psi\rangle = \sum_{m=0}^n e^{i\theta_m} \binom{n}{m}^{1/2} (1 - e^{-\zeta})^{m/2} e^{-\zeta(n-m)/2} |n-m\rangle_{\text{ph}} \otimes |m\rangle_{\text{ss}}, \quad (3.12)$$

which includes contributions from detector states $|m\rangle_{\text{ss}}$ with different numbers m of excited electrons—and hence, counted photons. We also note the generic, m -dependent phase $\exp(i\theta_m) \neq 1$ above, which is sensitive to microscopic details of the apparatus, may be time dependent and vary between experimental shots, and intuitively comes from phases other than $\pm i$ in H_k (3.7). Although the phases $\{\theta_m\}$ are not important in any single experiment shot or for detecting photons from a single mode, they are quite relevant in the more realistic settings of many electromagnetic modes (which often have spatiotemporal structure) and the extraction of statistics (which requires multiple shots and is sensitive to coherences). Moreover, the phases $\{\theta_m\}$ provide a mechanism for *decoherence*, which leads to the appearance of wavefunction collapse as we describe in Sec. 5.3.

For the initial Fock state $|\psi_0\rangle_{\text{ph}} = |n\rangle$ (3.12), the probability to detect $k \leq n$ photons is

$$p_k = \langle \Psi | \mathbb{1} \otimes P_{\text{ss}}^{(k)} | \Psi \rangle = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{with} \quad p = 1 - e^{-\zeta}, \quad (3.13)$$

corresponding to a binomial distribution, where $p = 1 - e^{-\zeta}$ is the probability that an incident photon is successfully detected. For the initial Fock state $|n\rangle$, the expected number of counted

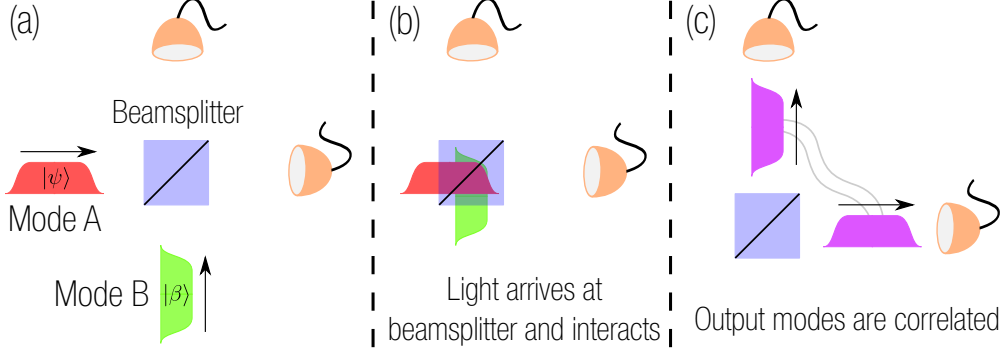


FIG. 2. Schematic depiction of homodyne detection. (a) An optical mode “A” of interest and a strong, coherent “B” mode (3.15) are sent towards a beam splitter. (b) The modes interact in the beam splitter. (c) The modes exit the beam splitter with coherent correlations and propagate towards photodetectors.

photons is np with variance $np(1-p)$. Hence, efficiency is greatest when the attenuation parameter ζ (3.9) is large. This also holds for more general initial states of the electromagnetic mode.

However, one can still effectively count photons even when $\zeta \lesssim 1$ is small (3.9). Essentially, when $p \ll 1$ (3.13) is small, the average number of excitations in the detector is $np \ll n$, with variance $np(1-p) \approx np$, so that the distribution is effectively Poisson, as one might expect in the context of particle detection. If np is large, the relative fluctuations with respect to the mean are small, with $\sim (np)^{-1/2} \ll 1$, so that if m excitations are observed, one can infer an occupation $n = m/p$ of the electromagnetic mode with high precision, provided that p is known.

Finally, we note that the simplified model of photon counting discussed above also applies to the measurement of *intensity* of narrow-band sources of light. For such sources, the intensity is simply proportional to photon number flux, even if the associated measurements rely on a physical mechanism different from the photodetection mechanism considered above [50]. We also note that Eq. 3.10 is the coherent representation of the statistics of single-mode attenuated fields originally derived in Refs. 59 and 60. In this context, the attenuation parameter ζ is replaced by $\kappa\tau$ from Ref. 59. In general, the mathematical form of the associated probabilities is agnostic to the specific detector model employed, as shown in Ref. 61, which also provides formulae for multi-time correlations. Other measurements of electromagnetic modes related to photon number include irradiance and excitance. For a discussion of how the model above is modified to account for the presence of many electromagnetic modes, we refer to the literature [62–64].

3.3. Balanced homodyne detection

Another important class of optical measurements relates to *interferometry* [65]. When combined with other ingredients, measurements of intensity or photon number (see Sec. 3.2) can be used to probe “quadratures” of light—i.e., a family of operators of the form

$$x_\phi \equiv \frac{1}{\sqrt{2}} \left(e^{-i\phi} a + e^{i\phi} a^\dagger \right) \quad (3.14)$$

which realize canonically conjugate operators for, e.g., $\phi = 0$ and $\phi = \pi/2$. Measuring such quadratures (3.14) can be achieved using homodyne detection [53–57, 66], where ϕ is the experimentally controllable “homodyne angle.” For convenience, we assume that the attenuation parameter $\zeta \gg 1$ (3.9) is large, so that the photodetectors (see Fig. 2) efficiently count photons [51].

Homodyne detection involves *two* electromagnetic modes of the *same* frequency, as schematically depicted in Fig. 2. The “A” mode represents the physical system of interest, prepared in an arbitrary initial state $|\psi\rangle_A$. The “B” mode is an auxiliary mode, prepared in the coherent state

$$|\beta\rangle_B = e^{-|\beta|^2/2} \sum_{n \in \mathbb{N}} \frac{\beta^n}{\sqrt{n!}} |n\rangle_B, \quad (3.15)$$

for complex $\beta \in \mathbb{C}$ with $|\beta| \gg 1$. This highly populated (i.e., bright) coherent state represents the field of a laser, which is commonly termed the “local oscillator” (LO) [54]. Its phase ($\beta = |\beta| e^{-i\phi}$) determines the quadrature x_ϕ (3.14) that is ultimately measured. For conceptual clarity, it is useful to regard both the A and B modes as propagating, localized wavepackets, as in Fig. 2(a). With this in mind, the steps involved in homodyne measurements are as follows (see Fig. 2):

1. The two electromagnetic modes are initially in the state $|\Psi_0\rangle = |\psi\rangle_A |\beta\rangle_B$. The modes propagate towards separate photodetectors [see Sec. 3.2 and Fig. 2(a)], each in the default initial state $|0\rangle_{\text{det}}$ with no excitations. The full initial state is $|\Psi_0\rangle \otimes |0, 0\rangle_{\text{ss}}$.
2. Before reaching the detectors, the A and B modes simultaneously pass through a beam splitter [see Fig. 2(b)], during which they evolve under the unitary [67, 68]

$$U_{\text{BS}} = \exp\left(\frac{\pi}{4}(a^\dagger b - ab^\dagger)\right), \quad (3.16)$$

where a and b are annihilation operators for the A and B modes, respectively. Applying U_{BS} (3.16) to the initial state $|\Psi_0\rangle$ produces a complicated state $|\Psi_1\rangle = U_{\text{BS}} |\Psi_0\rangle$.

3. The A and B modes exit the beam splitter and propagate toward their associated detector [see Fig. 2(c)], which counts the photons in that mode as in Sec. 3.2, i.e.,

$$\sum_{n_a, n_b \in \mathbb{N}} \left[\langle n_a, n_b | \Psi_1 \rangle \right] |n_a, n_b\rangle \otimes |0, 0\rangle_{\text{ss}} \mapsto \sum_{n_a, n_b \in \mathbb{N}} \left[\langle n_a, n_b | \Psi_1 \rangle \right] |0, 0\rangle \otimes |n_a, n_b\rangle_{\text{ss}}, \quad (3.17)$$

where $|n_a, n_b\rangle$ is a shorthand for $|n_a\rangle_A |n_b\rangle_B$. This process is described in Sec. 3.2 and Fig. 1.

4. We express the final state in terms of the sum $N = n_a + n_b$ and difference $D = (n_a - n_b)/2$ of the counts of the two detectors, so that the final state (3.17) is

$$|\Psi_f\rangle = \sum_{N=0}^{\infty} \sum_{D=-N/2}^{N/2} \langle N, D | U_{\text{BS}} | \Psi_0 \rangle |0\rangle_A |0\rangle_B \otimes |N, D\rangle_{\text{ss}}. \quad (3.18)$$

In Eq. 3.18, we introduced a new basis $|N, D\rangle$ for the physical and auxiliary modes, given by

$$|N, D\rangle = \frac{1}{\sqrt{(N/2 + D)!}} \frac{1}{\sqrt{(N/2 - D)!}} (a^\dagger)^{N/2+D} (b^\dagger)^{N/2-D} |0\rangle_A |0\rangle_B, \quad (3.19)$$

and all that remains is to calculate the matrix element $\langle N, D | U_{\text{BS}} | \Psi_0 \rangle$. As we show explicitly in App. C, in the limit $|\beta| \rightarrow \infty$ with $\langle \psi | a^\dagger a | \psi \rangle_A \ll |\beta|$ [66], this matrix element takes the form [52],

$$\langle N, D | U_{\text{BS}} | \Psi_0 \rangle = \frac{e^{iN\phi}}{\pi^{1/4} |\beta|} \exp\left[-\frac{(N - |\beta|^2)^2}{4|\beta|^2}\right] \psi_\phi(D\sqrt{2}/|\beta|), \quad (3.20)$$

where $\psi_\phi(x) = \langle x_\phi | \psi \rangle_A$ is the “quadrature wavefunction”—i.e., the overlap of the state $|\psi\rangle \in \ell^2(\mathbb{N})$ with the eigenbasis of the *quadrature* operator x_ϕ (3.14), instead of the usual $x = x_0$, i.e.,

$$\psi_\phi(x) = \langle x_\phi | \psi \rangle = \sum_{n=0}^{\infty} \frac{e^{-in\phi} e^{-x^2/2}}{\pi^{1/4} 2^{n/2} \sqrt{n!}} H_n(x) \langle n | \psi \rangle, \quad (3.21)$$

where $H_n(x)$ is the n th (physicist’s) Hermite polynomial. Since ϕ is defined by the phase of the coherent state (i.e., $\beta = |\beta| e^{-i\phi}$), the final state $|\Psi_f\rangle$ (3.18) can be written

$$|\Psi_f\rangle = \sum_{N=0}^{\infty} \sum_{D=-\infty}^{\infty} \frac{e^{iN\phi}}{\pi^{1/4} |\beta|} \exp\left[-\frac{(N - |\beta|^2)^2}{4|\beta|^2}\right] \psi_\phi(D\sqrt{2}/|\beta|) |0, 0\rangle \otimes |N, D\rangle_{ss}, \quad (3.22)$$

where we extend the sum over D from $\pm N/2$ to $\pm\infty$ by assuming that the quadrature wavefunction $\psi_\phi(D\sqrt{2}/|\beta|)$ vanishes for $D > N/2$. As a result, $|\Psi_f\rangle$ (3.22) is *separable* with respect to N and D .

Several remarks about the homodyne-detection procedure described above are in order:

- The total detector excitation level N is uncorrelated with the difference in excitations D , since $|\Psi_f\rangle$ (3.22) is a product of a function of N and a function of D , where

$$p(N) = \mathcal{N}(|\beta|^2, |\beta|^2) \quad \text{and} \quad p(D) = \left| \psi_\phi\left(D\sqrt{2}/|\beta|\right) \right|^2. \quad (3.23)$$

- The N -dependent part of $|\Psi_f\rangle$ (3.22) reflects a very bright coherent state, where $p(N)$ (3.23) realizes a normal (Gaussian) distribution whose mean and variance are both equal to $|\beta|^2$. Importantly, it encodes no information about the initial physical state $|\psi\rangle$.
- The coherences between different N sectors are extremely sensitive to the phase ϕ of the LO [52], which also appears in the factor $e^{iN\phi}$ in $|\Psi_f\rangle$ (3.22). Measuring coherences between macroscopically different values of N thus requires very precise control (e.g., with sensitivity $1/N$) of the phase ϕ of the LO, or else the coherences wash out upon averaging over many measurements, which leads to an effective source of decoherence (see also Sec. 5.3).
- The excitation difference D samples the quadrature wavefunction ψ_ϕ in $|\Psi_f\rangle$ (3.22). The quadrature basis x_ϕ (3.14) is determined by the phase ϕ of the LO (3.15), and the quadrature wavefunction $\psi_\phi(\cdot)$ is effectively probed at points separated by the spacing $\delta x_\phi = \sqrt{2}/|\beta|$, which vanishes when $|\beta|$ is large so that one resolves a continuum of values of $\psi_\phi(\cdot)$.
- The homodyne detection scheme described above operates in a “balanced” configuration, where the specific beam-splitter unitary U_{BS} (3.16) transmits (and reflects) 50% of the light in each mode. One can also use unbalanced configurations of the beam splitter, where the ratio between the reflection and transmission coefficients is not 1 : 1. Such setups can also be used to extract information about the state of electromagnetic modes [69–71].

Within this framework, each experimental shot results in a single “observed” value of D , sampled from the probability distribution $p(D) = |\psi_\phi|^2$ (3.23). Repeating the experiment with the same value of the homodyne angle ϕ (3.14) samples $p(D)$ with approximate spacing $\delta x_\phi = \sqrt{2}/|\beta|$, allowing for the reconstruction of $p(D)$ with sensitivity δx_ϕ . Moreover, repeating the experiment with *different* values of ϕ allows for the recovery of the full quantum state ψ_ϕ of mode A [57, 72].

In practice, extracting a value of D from an experimental shot requires integrating the instantaneous outputs of the A and B detectors over time, and computing sums and differences. The temporal nature of such interferometric measurement processes is especially pronounced in the

related context of *heterodyne detection*, in which the frequencies ω_A and ω_B are different. As a result, the quadrature (3.14) being probed changes as a function of time—i.e., the quadrature angle is $\phi(t) = (\omega_A - \omega_B)t$. Moreover, extracting information from this temporal distribution requires multiplying the time-dependent output of the detectors by a sinusoidal signal that oscillates at the beat frequency $\omega_A - \omega_B$ *before* integrating. Since quadratures (3.14) for different values of ϕ do not commute with one another, accommodating this technique into the Stinespring formalism—and particularly, extracting the *minimal* Stinespring form—requires more careful consideration of the time-dependent processes described above, which we defer to future work. Schematically, however, heterodyne detection can be viewed as effectively measuring the non-Hermitian annihilation operator a , whose coherent eigenstates are not mutually orthogonal and overcomplete [73].

4. QUBIT MEASUREMENTS

We now consider the measurements of various implementations of qubits, which involve the examples of photon measurements considered in Sec. 3. Although we restrict to two-level qubits, the ideas herein may also extend to multi-level qudits. The two levels of the qubit correspond to the computational (Z) basis states $|0\rangle$ and $|1\rangle$, where $Z|n\rangle = (-1)^n|n\rangle$. These states are generally associated with the internal energy levels of, e.g., atoms, ions, and superconducting circuits, where the interpretation varies depending on the underlying physical system. In atoms and ions, the levels correspond to distinct orbital and spin configurations of valence electrons; in superconducting circuits, the states describe different configurations of the superconducting current, phases, or voltages; in nitrogen-vacancy (NV) centers, the states correspond configurations of a defect and surrounding electrons; and in quantum dots, the configurations correspond to the presence of electrons in the potential wells or the spins of those electrons.

Typically, one seeks to measure the computational (Z) basis state of the qubit. Other single-qubit operators can be measured by first applying a unitary change of basis—e.g., applying the Hadamard gate $H = (X + Z)/\sqrt{2}$ and measuring Z is equivalent to measuring X . Certain multi-qubit operators can also be measured by applying entangling unitaries and then measuring Z . In contrast to the photon measurements of Sec. 3, qubit measurements are generically projective—as opposed to destructive—so that the postmeasurement state is an eigenstate of the measured operator.

In the remainder, we consider two types of measurements that are common across many different experimental systems: fluorescence state detection and dispersive readout. Since they ultimately rely on detection of light, we frequently invoke the results of Sec. 3. While other detection mechanisms for qubits exist, they generally seem to involve particle detection in some form.

4.1. Fluorescence measurement

Detecting fluorescent photons is a common means of measuring the Z -basis states of qubits realized in trapped ions [74], neutral atoms [75–78], NV centers [79, 80], and more. We first introduce a minimal model of fluorescence state detection for trapped-ion qubits, which is based on the “electron shelving” technique [74, 81]. We then explain how to adapt this simple model to fluorescence state detection of other experimental realizations of qubits.

The minimal model involves three atomic energy levels: the ground state $|g\rangle = |0\rangle$ and excited state $|e\rangle = |1\rangle$ that define the computational states of the qubit, and an *auxiliary* level $|a\rangle$, as depicted in Fig. 3. The state $|a\rangle$ is generally unstable (i.e., short lived, with a typical lifetime of a few nanoseconds). Fluorescence state detection operates as follows (see also Fig. 3). Incident light of the appropriate frequency $\omega_{ag} = E_a - E_g$ preferentially excites an atom from the state $|g\rangle = |0\rangle$ to the state $|a\rangle$, while leaving the state $|e\rangle = |1\rangle$ unaltered; the excited atom in the unstable state

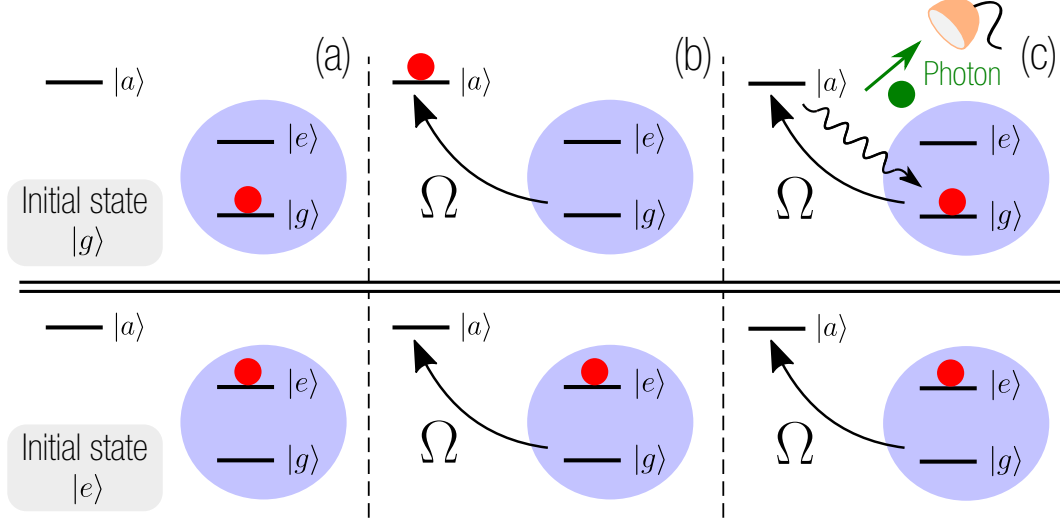


FIG. 3. State detection via fluorescent photons. (a) The top (bottom) row corresponds to the initial physical state $|g\rangle$ ($|e\rangle$). (b) Shining light on the system preferentially excites the state $|g\rangle$ to $|a\rangle$. (c) Relaxation via spontaneous emission of a photon that can be detected. This process is repeated many times.

$|a\rangle$ fluoresces, spontaneously emitting a photon as it relaxes back to the state $|g\rangle$. The process is then repeated, typically in a regime of high saturation (i.e., high intensity of incident light). The physical state is inferred from the presence ($|g\rangle$) or absence ($|e\rangle$) of emitted photons [74–78, 81].

This process is modeled using a nonminimal Stinespring Hilbert space $\mathcal{H}_{ss} = \mathcal{H}_{em} \otimes \mathcal{H}_{det}$, where $\mathcal{H}_{em}$ corresponds to the electromagnetic mode containing emitted (fluorescent) photons and $\mathcal{H}_{det}$ is the state of the photodetector. The electromagnetic mode incident on the atom is *not* included in $\mathcal{H}_{em}$; these photons merely induce the atomic transition $|g\rangle \rightarrow |a\rangle$. The state of the detector reflects the number of photons absorbed, as described in Sec. 3.2. The minimal representation (2.23) corresponds to associating $n > 0$ detected photons with $|g\rangle$ and $n = 0$ with $|e\rangle$.

The state of the full system (including the electromagnetic mode and detector) is initially

$$|\Psi_0\rangle = (c_g|g\rangle + c_e|e\rangle)_{ph} \otimes |0\rangle_{em}|0\rangle_{det}, \quad (4.1)$$

i.e., an arbitrary superposition of the qubit states $|0\rangle$ and $|1\rangle$, an unoccupied external electromagnetic field, and the default (unexcited) state of the photodetector. Fluorescence measurement begins with shining light on the atom, described by time evolution under the Hamiltonian

$$H = \frac{\omega_{ag}}{2}(|a\rangle\langle a| - |g\rangle\langle g|) + \frac{\Omega}{2}(\sigma_{ga}e^{i\omega_{ag}t} + \sigma_{ag}e^{-i\omega_{ag}t}) + H_{em} + H_{int}, \quad (4.2)$$

where $\sigma_{ga} = \sigma_{ag}^\dagger = |g\rangle\langle a|$ and Ω is the Rabi frequency of the driving light, which is resonant only with the atomic transition $|g\rangle \leftrightarrow |a\rangle$, H_{em} describes the evolution of the electromagnetic field, and H_{int} describes the interaction between the field and atom, including spontaneous emission of photons via the Wigner-Weisskopf mechanism. We do not write H_{int} explicitly, as it is quite complicated; we also stress that $|e\rangle$ does not appear in Eq. 4.2.

Physically, the atom is repeatedly excited from $|g\rangle$ to $|a\rangle$ by the incident light via the Rabi term in H with coefficient Ω in (4.2) [see Fig. 3(b)]. The interaction term H_{int} (4.2) captures the fluorescence process, where the atomic transition $|a\rangle$ to $|g\rangle$ is accompanied by a spontaneously emitted photon [see Fig. 3(c)]. The emitted photons evolve under H_{em} (4.2) until they are counted by the photodetector, as described in Sec. 3.2; the repeated excitation and emission increases the

chance of detection. Supposing that this process results in n photons, the intermediate state is

$$|\Psi_1\rangle = c_g|g\rangle_{\text{ph}} \otimes |n\rangle_{\text{em}}|0\rangle_{\text{det}} + c_e|e\rangle_{\text{ph}} \otimes |0\rangle_{\text{em}}|0\rangle_{\text{det}}, \quad (4.3)$$

where none of the photons have been absorbed by the photodetector [see Fig. 3(c)] for the purpose of illustration. The photons are then counted as described in Sec. 3.2, leading to a final state

$$|\Psi_f\rangle = c_g|g\rangle_{\text{ph}} \otimes |0\rangle_{\text{em}}|n\rangle_{\text{det}} + c_e|e\rangle_{\text{ph}} \otimes |0\rangle_{\text{em}}|0\rangle_{\text{det}}, \quad (4.4)$$

of the full system and detector, where we associate $n > 0$ counted photons with outcome g and $n = 0$ counted photons with outcome e (see also Fig. 3). In practice, however, the counting of emitted photons and reexcitation of the atom are concurrent, and not all emitted photons are necessarily counted. Although this may result in more complicated states than appear in Eqs. 4.3 and 4.4, we always associate $n > 0$ counted photons with $|g\rangle$ and $n = 0$ counted photons with $|e\rangle$.

The foregoing procedure is commonly referred to as “fluorescence state detection” [82] because it detects the Z -basis state. This is equivalent to measuring $Z = |g\rangle\langle g| - |e\rangle\langle e|$ itself. Importantly, associating $|n = 0\rangle_{\text{det}}$ with the outcome $|1\rangle$ (e) and *any* state $|n > 0\rangle_{\text{det}}$ with the outcome $|0\rangle$ (g) defines a binning procedure that leads to the minimal Stinespring unitary (2.23),

$$\mathcal{U}_{[Z]} = |g\rangle\langle g|_{\text{ph}} \otimes \mathbb{1}_{\text{ss}} + |e\rangle\langle e|_{\text{ph}} \otimes X_{\text{ss}}, \quad (4.5)$$

which acts on the initial state $|\psi\rangle_{\text{ph}} \otimes |e\rangle_{\text{ss}}$, where the default is e because it is associated with no change to the state of the detector; this is the general interpretation of the default state on $\mathcal{H}_{\text{ss}}$.

As with photon counting in Sec. 3.2, successful fluorescence measurement is possible even if the photodetector only counts emitted photons with probability $p < 1$ (3.13). Intuitively, one should then shine a large number of photons on the atom, in which case the postmeasurement state is

$$\begin{aligned} |\Psi_f(p)\rangle = & c_g|g\rangle \otimes \sum_{m=1}^n \left[\binom{n}{m} p^m (1-p)^{n-m} \right]^{1/2} |n-m\rangle_{\text{em}} |m\rangle_{\text{det}} \\ & + [c_e|e\rangle \otimes |0\rangle_{\text{em}} + c_g(1-p)^{n/2} |g\rangle \otimes |n\rangle_{\text{em}}] |0\rangle_{\text{det}}, \end{aligned} \quad (4.6)$$

where the two terms above correspond to $m \geq 1$ photons detected and $m = 0$ photons detected, respectively. For any $p > 0$, if $m > 0$ photons are detected, the atom is guaranteed to be in the state $|g\rangle$, provided that sources of dark counts (i.e., false positives) are mitigated. However, for $p < 1$, there is a nonzero probability to detect $m = 0$ photons even if the atom is in the excitable state $|g\rangle$. If p is the probability that a photon shone on the atom induces fluorescence *and* the spontaneously emitted photon is absorbed by the detector, then the probability of such a “false negative” is $|c_g|^2 (1-p)^{n/2}$, where n is the number of photons shone on the atom.

Hence, shining a very large number of photons on the atom exponentially reduces the probability of a false negative. In this limit, the final state can be represented schematically as

$$|\Psi_f\rangle = c_g|g\rangle_{\text{ph}} \otimes |\text{clicks}\rangle_{\text{ss}} + c_e|e\rangle_{\text{ph}} \otimes |\text{no clicks}\rangle_{\text{ss}}, \quad (4.7)$$

which we note is of the minimal Stinespring form now that we have binned “clicks” ($m > 0$) versus “no clicks” ($m = 0$). However, scattering large numbers of photons leads to unwanted heating so that, in practice, one must optimize the number of photons shone on the atom.

We now discuss some important practical considerations involved in the experimental implementation of the foregoing simplified model of fluorescence measurements of Z in atomic systems, as well as how the technique is commonly extended to other realizations of qubits.

- The fluorescence measurement process described above is most commonly used to determine the energy levels of, e.g., trapped ions [83–85] and neutral atoms [86–88]. With neutral atoms, it is often preferable to induce the transition $|g\rangle \rightarrow |a\rangle$ with off-resonant light to avoid, e.g., heating-induced losses and state leakage [89]. Alternatively, cavities can be used to enhance the collection efficiency [90] and ensure that light is preferentially emitted in particular directions to facilitate the collection of outgoing photons [91].
- Fluorescence measurements are also used to determine the occupancy of “sites” in optical lattices [76, 77, 92]. The states $|g\rangle$, $|a\rangle$ and $|e\rangle$ above are replaced by $|1, g\rangle$, $|1, a\rangle$ and $|0\rangle$, where $|1, g/a\rangle$ indicates the presence of an atom in the state g/a and $|0\rangle$ indicates the absence of an atom or an atom in a state other than g or a . We also note that conservation of particle number requires that superpositions of $|0\rangle$ and $|1, g/a\rangle$ on a given “site” must be correlated with mixed occupations of other sites [93]. This necessitates a description of the measurement process involving Stinespring registers for each site; in many experimental implementations, the imaging technique only discriminates between even and odd occupations of several optical lattice sites, as atoms are typically lost in pairs due to light-assisted collisions [94].
- Fluorescence measurements can also be used as an indirect probe of which “sites” in an optical lattice are occupied by an atom in the state g . This is particularly common in tweezer arrays, where one preferentially expels all atoms in the state e (e.g., via resonant heating [95, 96] or due to antitrapping [97]). One then uses fluorescence measurement to obtain a “snapshot” of the locations of the remaining atoms, which are in the state g . This requires the ability to resolve light emitted from different lattice sites. We also note that this process is *destructive*, as the ejected atoms are irrevocably lost, but is useful in numerous settings.
- Fluorescence state detection also applies to NV centers, whose various internal states involve configurations of, e.g., electronic spins in the vicinity of the defect [79, 80]. In contrast to the atomic implementation above, fluorescence measurement of NV centers require *two* auxiliary internal states $|a\rangle$ and $|a'\rangle$. The light incident on the NV center generically induces transitions $|g\rangle \rightarrow |a\rangle$ and $|e\rangle \rightarrow |a'\rangle$, but the various decay pathways out of $|a\rangle$ versus $|a'\rangle$ are different. The system can return to both of the states $|g\rangle$ and $|e\rangle$ via radiative processes, and can return to the state $|g\rangle$ via nonradiative decay (via intersystem crossing). Importantly, these nonradiative processes occur with different rates, which has two consequences: (i) at the end of the measurement the NV center realizes the state $|g\rangle$ regardless of its initial state and (ii) the number of scattered photons depends on whether the initial state was $|g\rangle$ or $|e\rangle$. The final state following the measurement is given schematically by [80]

$$|\psi\rangle = [c_g |g\rangle + c_e |e\rangle] \otimes |0\rangle_{\text{em}} \mapsto |g\rangle \otimes [c_g |n_g\rangle_{\text{em}} + c_e |n_e\rangle_{\text{em}}], \quad (4.8)$$

where $|n_{g,e}\rangle$ is a placeholder for some complicated state of the electromagnetic field that, on average, has $n_{e,g}$ photons. It is typically assumed to be coherent to account for Poisson statistics [80]. The outcomes corresponding to $|e\rangle$ versus $|g\rangle$ are more subtle to differentiate than “clicks” versus “no clicks,” since the associated numbers of emitted photons are very small. However, performing the experiment many times increases the total numbers of collected photons from the states $|g\rangle$ and $|e\rangle$; it also increases their *difference*, so that one can reliably distinguish the two states. Note that $|\psi\rangle$ (4.8) has the Stinespring form of a destructive measurement (2.33).

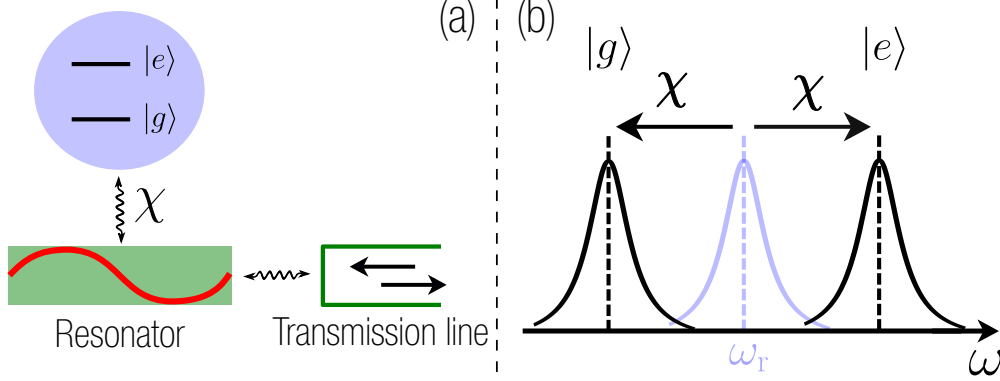


FIG. 4. Left: Schematic of a superconducting qubit interacting with a resonator that is coupled to a transmission line. Right: Bare cavity frequency (ω_r) and the qubit state-dependent frequency shifts (blue).

4.2. Dispersive measurement

Here we describe dispersive state readout [98]—a common technique for determining the Z -basis state of superconducting qubits and atoms in cavities (microwave and optical). For concreteness, we focus on a single superconducting qubit, whose two levels generally represent physically distinct configurations of charges, phases, and/or fluxes in a superconducting circuit. This state can be measured indirectly via controllable interactions with auxiliary systems, as depicted schematically in Fig. 4. More concretely, one sends a pulse of light through a transmission line to probe a resonator (e.g., a cavity) that hosts confined electromagnetic modes. The interaction of the resonator modes with the qubit modify the pulse in the transmission line (see Fig. 4), from which the qubit's state is inferred. A minimal description consists of the qubit (with states g and e), a single resonator mode (with photon annihilation operator r), and the transmission line (whose continuum of propagating electromagnetic modes is represented schematically by Fock states $|n\rangle_{\text{out}}$). As in Sec. 3, we picture these outgoing modes as being mildly localized and propagating towards a detector.

The measurement protocol involves evolution under the qubit-resonator Hamiltonian

$$H_{\text{QR}} = \omega_a \sigma^z + \omega_r r^\dagger r + g(r^\dagger \sigma^- + r \sigma^+), \quad (4.9)$$

where ω_a is the frequency of transitions between the states $|e\rangle$ and $|g\rangle$ of the qubit, ω_r is the frequency of the resonator, and g controls the strength of a Jaynes-Cummings interaction [99], which models coherent photon absorption and emission by the qubit. In the *dispersive* regime—where the large detuning $|\omega_a - \omega_r| \gg g$ ensures that the atom quickly emits any absorbed photons—we perturb about $g = 0$ to good approximation, recovering the Hamiltonian

$$H_{\text{QR}} = \omega_a \sigma^z + \omega_r r^\dagger r + \chi(r^\dagger r) \sigma^z, \quad (4.10)$$

where $\chi = g^2/(\omega_r - \omega_a)$, and this term describes a shift in the resonator frequency ω_r , whose sign depends on the qubit's Z -basis state [see Fig. 4(b)]. Dispersive readout proceeds as follows:

1. The qubit of interest is prepared in an arbitrary superposition $|\psi\rangle = c_g|g\rangle + c_e|e\rangle$ and the field modes (resonator and transmission line) are initially empty. The dilated initial state is

$$|\Psi_0\rangle = (c_g|g\rangle + c_e|e\rangle) \otimes |0\rangle_r |0\rangle_{\text{out}}. \quad (4.11)$$

2. A mode with the same frequency ω_r as the *bare* resonator is sent to the resonator via the transmission line. The outgoing modes realize a coherent state $|\alpha\rangle$ (3.15), and the average

number of photons $n \sim |\alpha|^2$ sent towards the resonator grows linearly in time in the late-time steady-state regime of the electromagnetic modes. The state remains separable, with

$$|\Psi_1\rangle = |\psi\rangle \otimes |0\rangle_r |\alpha\rangle_{\text{out}}. \quad (4.12)$$

3. Photons sent through the transmission line enter the resonator, interact with the qubit, and escape back through the line. The resonator responds to the incoming light as a damped-driven harmonic oscillator (damped by the modes in the transmission line). Crucially, this response depends on the resonator frequency, which in turn depends on the state of the qubit via χ (4.10). This process (with the foregoing assumptions) is captured by:

$$|g\rangle \otimes |0\rangle_r |\alpha\rangle_{\text{out}} \mapsto |g\rangle \otimes |\rho e^{+i\theta}\rangle_r |0\rangle_{\text{out}} \mapsto |g\rangle \otimes |0\rangle_r |\alpha e^{+2i\theta}\rangle_{\text{out}} \quad (4.13a)$$

$$|e\rangle \otimes |0\rangle_r |\alpha\rangle_{\text{out}} \mapsto |e\rangle \otimes |\rho e^{-i\theta}\rangle_r |0\rangle_{\text{out}} \mapsto |e\rangle \otimes |0\rangle_r |\alpha e^{-2i\theta}\rangle_{\text{out}}, \quad (4.13b)$$

where $|\rho e^{\pm i\theta}\rangle_r$ and $|\alpha e^{\pm 2i\theta}\rangle_{\text{out}}$ are coherent states (3.15). The final state of the system is

$$|\Psi_f\rangle = c_g |g\rangle \otimes |0\rangle_r |\alpha e^{2i\theta}\rangle_{\text{out}} + c_e |e\rangle \otimes |0\rangle_r |\alpha e^{-2i\theta}\rangle_{\text{out}}. \quad (4.14)$$

In this schematic description, the light enters the cavity and establishes an intracavity field $\rho e^{\pm i\theta}$ whose phase depends on the state of the qubit ($+i\theta$ for $|g\rangle$ and $-i\theta$ for $|e\rangle$). The coherent light then escapes the resonator via the transmission line, so that its phase can be recorded. We comment that, in practice, all of these processes occur at the same time and the device is operated in steady-state conditions after transients have died off.

4. The phase of the returning light is extracted via homodyne detection of the momentum quadrature $p = (a - a^\dagger)/i\sqrt{2} = x_{\phi=\pi/2}$ (3.15). Working in the eigenbasis of p ,

$$|\alpha e^{\pm i\theta}\rangle = \pi^{-1/4} \int_{-\infty}^{\infty} dp e^{-i\sqrt{2}\alpha p \cos \theta} e^{-(p \mp \sqrt{2}\alpha \sin \theta)^2/2} |p\rangle \equiv \int_{-\infty}^{\infty} dp \psi_{\pm}(p) |p\rangle, \quad (4.15)$$

and, since the distribution of the output momentum for the initial qubit states $|g\rangle$ and $|e\rangle$ are centered at $p = \pm\sqrt{2}\alpha \sin \theta$, respectively, we associate the detection of $|p > 0\rangle$ with $|g\rangle$ and of $|p < 0\rangle$ with $|e\rangle$. This procedure is not perfect: The probability to detect $p < 0$ ($p > 0$) when the initial state of the qubit is g (e) is given by

$$p_{\text{error}} = \int_0^{\infty} dp |\psi_-(p)|^2 = \int_{-\infty}^0 dp |\psi_+(p)|^2 = \frac{1}{2} \text{erfc}(\alpha \sqrt{2} \sin \theta), \quad (4.16)$$

which saturates to $1/2$ when $\alpha \sin \theta \rightarrow 0$, and vanishes as $\exp(-\alpha^2 \sin^2 \theta)/\alpha \sin \theta$ in the limit $\alpha \sin \theta \rightarrow \infty$. In this limit of many outgoing photons, the final state takes the form

$$|\Psi_f\rangle = c_g |g\rangle \otimes |0\rangle_r \left(\int_0^{\infty} dp \psi_+(p) |p\rangle_{\text{out}} \right) + c_e |e\rangle \otimes |0\rangle_r \left(\int_{-\infty}^0 dp \psi_-(p) |p\rangle_{\text{out}} \right), \quad (4.17)$$

which requires that θ is within $(0, \pi/2)$, or else one cannot discriminate between these outcomes; however, this is generally the case in experiment. The outgoing light is then measured via homodyne detection, so that correlations between the qubit and the electromagnetic mode (in the transmission line) are transferred into correlations between the qubit and the

differential intensity measurements of the final detectors, as in Eq. 3.22. In actual experimental implementations, the phase θ in Eq. 4.13 is measured in a steady-state configuration in which there is a fixed *flux* of photons (i.e., intensity) emanating from the resonator. By waiting longer, transients die off and $|\alpha|^2$ grows larger, linearly in time, reducing the probability of a readout error. However, separate experimental factors may degrade the measurement process over time, meaning there exists an optimal duration for the dispersive-readout process.

We now describe the application of this detection scheme in various experimental settings:

- Dispersive readout as described above is a common measurement approach for superconducting qubits [98] and detecting the presence of absence of atoms in cavities [100].
- The same scheme can also be used to measure the collective “inversion” of many atoms—i.e., the difference between the number of atoms in the excited versus ground state—in cavities (both optical and microwave) [101]. Instead of measuring σ^z for a single atom, one measures $S_z = \sum_j \sigma_j^z/2$ for an ensemble of N atoms. The role of the resonator is played by the cavity itself and that of the transmission-line modes is played by light leaking through the cavity mirrors, which can be measured using homodyne (or heterodyne) detection.

In the “dispersive regime,” the atom-cavity interaction is given by Eq. 4.10, with σ^z replaced by $2S_z$. The detection scheme proceeds as described above. Note that, in this collective regime, there are $N + 1$ possible values S_z instead of two. Ideally, the procedure collects enough photons to enable almost perfect discrimination between each value of inversion, yielding a postmeasurement state in Stinespring form (4.17). Schematically, if $|m\rangle$ represents the different eigenstates of the inversion S_z , the dispersive-readout process takes the form

$$|\Psi_0\rangle = \sum_{m=-N/2}^{N/2} c_m |m\rangle \otimes |0\rangle_r |0\rangle_{\text{out}} \mapsto |\Psi_f\rangle = \sum_{m=-N/2}^{N/2} c_m |m\rangle \otimes |0\rangle_r |\alpha e^{im\theta}\rangle_{\text{out}}, \quad (4.18)$$

where the coherent states $|\alpha e^{im\theta}\rangle$ are nearly orthogonal for different m -dependent phases $m\theta$. This is often not possible due to technical limitations, but the information gained by the measurement can still be used to constrain the most likely values of S_z , in many cases creating entanglement between atoms in the ensemble [101–104].

The physical mechanism underlying this collective S_z measurement—namely, the state-dependent shift of the cavity (resonator) frequency captured by Eq. 4.10—is, strictly speaking, not restricted to the dispersive regime. It also occurs when the atoms and cavity are close to (or on) resonance—i.e., $\omega_a = \omega_r$ in Eq. 4.9—though the functional relationship between the atomic state and the shift in cavity frequency is different. In contrast to the dispersive scheme, three levels are typically used, where the cavity is resonant with the transition $|g\rangle \leftrightarrow |a\rangle$, rather than $|e\rangle \leftrightarrow |g\rangle$, for some auxiliary state $|a\rangle$. The objective of a measurement in this regime is not to infer the inversion $S_z = N_e - N_g$, but simply the number N_g of atoms in $|g\rangle$ [102, 103]. The inversion can be determined using a similar, complementary measurement of N_e . These resonant measurements can also be used in the single-atom regime [105].

- The dispersive response of atoms can also be used directly with atomic clouds to infer their average atomic state [106], or for spatial imaging, whereby a probe beam of light acquires spatially dependent phase shifts as it interacts with the atoms in the cloud [107–109]. We defer a Stinespring description of this procedure to the future, since it involves the spatial distribution of the electromagnetic field, which we have ignored thus far.

5. USING THE UNITARY MEASUREMENT FORMALISM

Having discussed several prominent experimental implementations of projective and destructive measurements and their connection to the dilated unitary formalism presented in Sec. 2, we now discuss how to *use* this formalism in practice and its various advantages, implications, and properties.

5.1. Measurement outcomes and the Born rule

We begin by showing that the standard concepts associated with such measurements recover in the Stinespring representation of Sec. 2. In particular, we show how the standard Born rule may be evaluated either before or after the measurement, and is a natural consequence of Gleason’s theorem [110]. We also consider the extraction of expectation values. As noted in Sec. 2, the projective measurement of an observable $\mathcal{O}$ (2.3) is always associated with a quantum *channel*

$$\Phi(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger \right) = \rho_{\text{av}}, \quad (2.15)$$

where the *measurement unitary* $\mathcal{U}$ acts on $\mathcal{H}_{\text{dil}} = \mathcal{H} \otimes \mathcal{H}_{\text{ss}}$, with $\mathcal{H}_{\text{ss}}$ the state space of the measurement apparatus and $|i\rangle_{\text{ss}}$ the default initial state of the apparatus. The channel Φ is CPTP, and ρ_{av} is the outcome-averaged postmeasurement density matrix.

Related to Φ (2.15) is the selective operation [20] corresponding to the quantum *channel*

$$\Phi_m(\rho) = \text{tr}_{\text{ss}} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger P_{\text{ss}}^{(m)} \right) = p_m \rho_m, \quad (2.16)$$

which is a CP map describing a projective measurement resulting in the particular outcome m , where $P_{\text{ss}}^{(m)}$ projects onto the state(s) of the apparatus associated with outcome m , p_m is the probability of observing outcome m , and ρ_m is the “collapsed” postmeasurement density matrix [20].

The foregoing expressions for Φ (2.15) and Φ_m (2.16) only apply to projective measurements. The destructive measurements discussed in Sec. 2.3.5 generally result in the *same* state of the physical system for *any* outcome m . Apart from the postmeasurement state, however, all of the formulae below apply equally to projective and destructive measurements.

The probability to observe outcome m (i.e., the eigenvalue O_m) is given by the *Born rule*

$$p_m = |\langle m|\psi\rangle|^2 \quad \longleftrightarrow \quad p_m = \text{tr} (P_m \rho), \quad (5.1)$$

where ρ is the premeasurement state of the system. The familiar expression on the left holds when $\rho = |\psi\rangle\langle\psi|$ is pure and the m th outcome is nondegenerate (2.3); the expression on the right is generic. Importantly, the Born rule (5.1) is generally an assumption of quantum mechanics—i.e., it is built into the axioms [1–8]. We note that the state ρ encodes a probability distribution, which one expects can be extracted from expectation values as in Eq. 5.1. The expectation value of an operator is a C^* -state—a positive linear functional. Positivity of the spectral projectors (2.1) then implies that $p_m \geq 0$, and because the projectors form a complete set (2.2), we have

$$0 \leq p_m = \text{tr} (P_m \rho) \leq \sum_{m=0}^{\mathcal{N}-1} \text{tr} (P_m \rho) = \text{tr} (\rho) = 1, \quad (5.2)$$

so that the probabilities p_m are positive and sum to one, as required. Moreover, Gleason’s theorem [110] establishes that, *a priori*, probabilities are always related to projectors via Eq. 5.2.

Using the Stinespring representation of measurements developed in Sec. 2.3, we can write

$$p_m = \text{tr} \left(\mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger P_{\text{ss}}^{(m)} \right) = \text{tr} [\Phi_m(\rho)] = p_m \text{tr} (\rho) = p_m, \quad (5.3)$$

i.e., p_m is the expectation value of the Stinespring projector $P_{\text{ss}}^{(m)}$ in the postmeasurement state

$$\varrho(t) = \mathcal{U} \rho \otimes |i\rangle\langle i|_{\text{ss}} \mathcal{U}^\dagger = \sum_{m=0}^{\mathcal{N}-1} p_m^{-1} \Phi_m(\rho) \otimes P_{\text{ss}}^{(m)}, \quad (5.4)$$

on $\mathcal{H}_{\text{dil}}$. As always, the Born probability corresponds to the expectation value of a projector (5.2), as guaranteed by Gleason's theorem [110]. The probability p_m may be extracted from the physical premeasurement state ρ_0 (5.8) via Eq. 5.1 or the dilated postmeasurement state $\varrho(t)$ (5.4) via Eq. 5.3. The latter has a physical interpretation as the expectation value (or probability) of finding the detector in a state corresponding to outcome m following evolution under $\mathcal{U}$ (2.17).

Following a projective measurement resulting in outcome m , the system's state is

$$\rho_m = p_m^{-1} \Phi_m(\rho) = \frac{P_m \rho P_m}{\text{tr}(P_m \rho)}, \quad (5.5)$$

provided that $p_m > 0$. The average postmeasurement state ρ_{av} is given by $\Phi(\rho)$ (2.15). However, in the dilated unitary representation, these merely correspond to particular limits of the *dilated* postmeasurement density matrix ϱ (5.4) of the system *and* apparatus. The average over all outcomes recovers upon taking the trace over the apparatus Hilbert space $\mathcal{H}_{\text{ss}}$ (2.15), giving the reduced postmeasurement density matrix of the physical system. The collapsed density matrix ρ_m (5.5) corresponds to projecting the apparatus onto the corresponding state $|m\rangle_{\text{ss}}$, and then taking the trace (2.16). Because the dilated unitary $\mathcal{U}$ (2.17)—or Hamiltonian H (2.34)—is both guaranteed by the axioms (see Sec. 2) and consistent with known experimental implementations (see Secs. 3 and 4), we adopt the perspective that ϱ is the correct postmeasurement density matrix, in which *all outcomes occur*. In Sec. 5.3, we describe why only one outcome is observed in practice.

Finally, the expectation value for measuring $\mathcal{O}$ in some state ρ can be written

$$\langle \mathcal{O} \rangle_\rho = \sum_{m=0}^{\mathcal{N}-1} p_m O_m = \sum_{m=0}^{\mathcal{N}-1} O_m \text{tr}(P_m \rho) = \text{tr}(\mathcal{O} \rho), \quad (5.6)$$

in terms of the premeasurement state ρ , or in terms of the postmeasurement state ϱ (on $\mathcal{H}_{\text{dil}}$) via

$$\langle \mathcal{O} \rangle_\rho = \sum_{m=0}^{\mathcal{N}-1} p_m O_m = \sum_{m=0}^{\mathcal{N}-1} \text{tr}_{\text{dil}} \left[\left(\mathbb{1}_{\text{ph}} \otimes O_m P_{\text{ss}}^{(m)} \right) \varrho \right], \quad (5.7)$$

which reproduces Eq. 5.6, as expected. The expectation value and Born probabilities are the same for projective and destructive measurements, and may be evaluated either before or after the measurement unitary $\mathcal{U}$ is applied using ρ or $\varrho(t)$, respectively. As a reminder, the fact that the Born probability takes the form of an expectation value of a projector is a consequence of Gleason's theorem [110], which extends without caveat to the dilated Hilbert space $\mathcal{H}_{\text{dil}}$ (2.8).

5.2. Application to measurement-based protocols

Perhaps the greatest advantage of the Stinespring formulation of projective and destructive measurements is the application to *adaptive* quantum protocols. These protocols involve nonlocal quantum operations based on the outcomes of prior measurements, and are known to be faster (and sometimes less error prone) than local unitary—or even Lindblad—dynamics [21, 111, 112]. However, such protocols are often cumbersome to describe in both the Kraus (2.11) or von Neumann (2.34)

representations, especially if one is interested calculations that require evolving operators—e.g., diagnosing locality, constraining quantum protocols, and identifying optimal strategies [21, 23].

The advantage of representing measurements unitarily on a dilated Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H}_{\text{ph}} \otimes \mathcal{H}_{\text{ss}}$ (2.8) is that the outcomes are reflected in explicit degrees of freedom in $\mathcal{H}_{\text{ss}}$, corresponding to the states of the detectors. In experiment, distinct detectors are not necessarily required, and one can simply reset and reuse the same detector for multiple measurements and “write down” the outcome. However, for the purpose of calculation, it is most convenient to encode the measurement outcomes in distinct Hilbert spaces. We may also consider protocols in which the decision of *whether* or not to perform a measurement (or what operator to measure) is conditioned on prior outcomes. One simply assigns Stinespring degrees of freedom to each *possible* measurement, and if a measurement is not performed, one simply does not use the corresponding Stinespring register.

Consider a quantum protocol $\mathcal{W}$ involving measurements of up to M observables $\{\mathcal{O}_j\}$, where all aspects of $\mathcal{W}$ at any time t may depend on *any* prior measurement outcomes. The physical system is initialized in an arbitrary state ρ_0 on $\mathcal{H}_{\text{ph}}$, and M Stinespring registers labeled j are initialized in the default state $|0\rangle\langle 0|_{\text{ss},j}$ of each apparatus, so that

$$\varrho(0) = \rho_0 \otimes |\mathbf{0}\rangle\langle \mathbf{0}|_{\text{ss}}, \quad (5.8)$$

is the dilated initial state. The physicist’s Stinespring theorem [20, 30, 31] ensures that $\mathcal{W}$ acts unitarily on $\mathcal{H}_{\text{dil}}$ (2.8), which may require enlarging $\mathcal{H}_{\text{ss}}$ to include “environmental” degrees of freedom. For simplicity, we consider protocols $\mathcal{W}$ involving only measurements and unitary time evolution on $\mathcal{H}_{\text{ph}}$, which may be arbitrarily conditioned on prior outcomes. The unitary evolution may be generated by a discrete quantum circuit or continuous evolution under some local Hamiltonian $H(t)$. In the latter case, one must either approximate the measurement unitary $\mathcal{U}$ as instantaneous or use the Hamiltonian formulation of measurements [21].

For convenience, suppose that $\mathcal{W}$ is realized by a unitary circuit on $\mathcal{H}_{\text{dil}}$ (2.8). We now consider the distinct types of dilated unitary “gates” in $\mathcal{W}$. Time evolution is realized by a unitary U acting only on $\mathcal{H}_{\text{ph}}$, while a projective measurement of an observable $\mathcal{O}_j$ (2.3) takes the form

$$\mathcal{U}_j = \sum_{n=0}^{\mathcal{N}_j-1} P_j^{(n)} \otimes X_{\text{ss},j}^n, \quad (5.9)$$

where $P_j^{(n)}$ projects onto the n th eigenspace of $\mathcal{O}_j$ (2.3) associated with Stinespring label j . When $\mathcal{N}_j \rightarrow \infty$, $\mathcal{U}_j$ (5.9) is replaced with a dilated Hamiltonian (2.34) or with a modified unitary (2.28). For destructive measurements, one modifies the physical part of Eq. 5.9 to match, e.g., Eq. 2.33.

The measurement of $\mathcal{O}_j$ (2.3) via $\mathcal{U}_j$ (5.9) entangles the system with the j th Stinespring register, which reflects the observed outcome. Importantly, we do *not* immediately trace over $\mathcal{H}_{\text{ss},j}$ to realize the quantum channel Φ (2.15) or the quantum operation Φ_m (2.16). Instead, we view $\mathcal{U}_j$ (5.9) as the time evolution of the system and the j th detector during the measurement of $\mathcal{O}_j$ (2.3), and defer *all* traces until the end of the *full* calculation. The relevant quantities are given by

$$\rho_{\text{av}} = \text{tr}_{\text{ss}} \left(\mathcal{W} \varrho(0) \mathcal{W}^\dagger \right) \quad (5.10a)$$

$$p_{\mathbf{n}} \rho_{\mathbf{n}} = \text{tr}_{\text{ss}} \left(\mathcal{W} \varrho(0) \mathcal{W}^\dagger P_{\text{ss}}^{(\mathbf{n})} \right), \quad (5.10b)$$

where $P_{\text{ss}}^{(\mathbf{n})} = |n_1, \dots, n_M\rangle\langle n_1, \dots, n_M|_{\text{ss}}$ projects (i.e., collapses) the dilated state onto a particular sequence $\mathbf{n} = \{n_1, \dots, n_M\}$ of measurement outcomes (5.10b), and taking the trace averages over all outcomes (5.10a). Importantly, these traces are only taken *after* $\mathcal{W}$ has been applied in its entirety, to allow for subsequent outcome-dependent operations. As a bookkeeping tool, one can

view the unitary description in Eq. 5.10 as an algebraic means of accounting for all Kraus operators—including outcome-dependent operations—in a convenient manner. However, we associate $\mathcal{W}$ with the *physical* time-evolution operator for the system *and* detectors throughout the protocol.

Because the trace is only taken at the end in Eq. 5.10, we can express a unitary *feedback* operation on the physical degrees of freedom in $\mathcal{H}_{\text{ph}}$ that is conditioned on the j th measurement outcome via

$$\mathcal{R}_j = \sum_{n=0}^{\mathcal{N}_j-1} U_n \otimes P_{\text{ss}}^{(n)}, \quad (5.11)$$

which applies the physical unitary U_n if the j th outcome was n , where the trivial case corresponds to $U_n = \mathbb{1}$ (i.e., do nothing). Such outcome-dependent unitaries are essential to measurement-based protocols, as without these operations, the outcome-averaged final state $\rho(t)$ is some maximally mixed state [21–23]. We note that Eq. 5.11 is easily extended to include physical unitaries that are conditioned on multiple outcomes—e.g., the parity of several qubit measurements, in the context of quantum teleportation [23]. One can also describe *conditional measurements* via

$$\mathcal{M}_{j,k} = \sum_{n=0}^{\mathcal{N}_j-1} \mathcal{U}_{[\mathcal{O}_n]} \otimes P_{\text{ss}}^{(n)}, \quad (5.12)$$

which realizes a projective measurement of an observable $\mathcal{O}_n$ if the j th measurement resulted in outcome n . The measurement unitary $\mathcal{U}_n$ acts on $\mathcal{H}_{\text{ss}}$ and the k th Stinespring register—in which the measurement outcome is encoded—conditioned on the outcome of the j th measurement. As with outcome-dependent feedback (5.11), one can also extend Eq. 5.12 to measurements conditioned on multiple outcomes, and the case in which no measurement is performed if the j th measurement resulted in outcome m is captured by replacing the n th measurement unitary with $\mathbb{1}$.

5.3. Decoherence and “wavefunction collapse”

For the dilated unitary $\mathcal{U}$ (5.9) to be considered *physical*, we must explain why only one outcome is *observed* in any experimental measurement. It is widely understood that quantum *decoherence*—the mechanism by which classical physics emerges in quantum systems—explains the appearance of wavefunction collapse [15]. Here, we combine recent results on quantum chaos and thermalization with the Stinespring representation of measurements outlined in Sec. 2.

Importantly, the physicist’s Stinespring theorem (2.51) implies a bookkeeping representation of measurements in which all outcomes occur *deterministically* on $\mathcal{H}_{\text{dil}}$ (2.8)—i.e., the dilated postmeasurement state $\varrho(t)$ (5.4) is a *coherent* superposition of states $\propto |\psi_m\rangle \otimes |m\rangle_{\text{ss}}$, where the state $|m\rangle$ of the apparatus indicates that the system is in the state $|\psi_m\rangle$. Although the *quantum* probability distributions encoded by such coherent superpositions may exhibit interference, this is not observed in any particular experimental instance of a measurement.

Instead, the fact that only one outcome is observed indicates that $\varrho(t)$ (5.4) encodes an *incoherent* (i.e., classical) probability distribution. For the ensemble of states encoded in the postmeasurement state $\varrho(t)$ (5.4) to appear classical, $\varrho(t)$ must realize a *mixed state*, i.e.,

$$\varrho_{\text{mix}} = \sum_{m=0}^{\mathcal{N}-1} p_m |\psi_m\rangle\langle\psi_m|, \quad (5.13)$$

where $\rho_m = |\psi_m\rangle\langle\psi_m|$ is pure and p_m is the classical probability for outcome m . While the coefficients of pure states correspond to coherent superpositions in a particular basis and reflect

quantum probabilities, the coefficients of the different pure states in a mixed state (5.13) reflect incoherent *classical* probabilities. Mixed states have always had the physical interpretation of a classical ensemble of pure quantum states ρ_m , only one of which is observed classically. A given experiment realizes a single pure state ρ_m , which is sampled from ϱ_{mix} with probability p_m .

It remains to identify a mechanism by which $\varrho(t)$ (5.4) realizes a mixed state ϱ_{mix} (5.13), i.e.,

$$\varrho(t) = \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho_0 P_n \otimes |m\rangle\langle n|_{\text{ss}} \mapsto \sum_{m=0}^{\mathcal{N}-1} P_m \rho_0 P_m \otimes P_{\text{ss}}^{(m)} = \sum_{m=0}^{\mathcal{N}-1} p_m \rho_m \otimes P_{\text{ss}}^{(m)}, \quad (5.14)$$

where we used Eqs. 5.8 and 5.9, and $\rho_m \otimes P_{\text{ss}}^{(m)}$ realize pure states. Understanding how the process in Eq. 5.14 comes about in actual experiments involves two key observations.

The first observation is that the state of the apparatus is classically observable and in thermal equilibrium with its environment. The eigenstate thermalization hypothesis (ETH) [113–115]—which has since been proven [116]—states that entanglement between a system and its environment is the mechanism for quantum thermalization. Because the apparatus is thermal, it is guaranteed to be highly (if not maximally) entangled with the environment. Studies of the chaotic quantum dynamics leading to thermalization have shown that the dynamics is well described by evolution under an ensemble of random matrices with the appropriate symmetries [22, 117–122].

The second observation is that the postmeasurement state of the apparatus is stable to perturbations. In other words, the chaotic quantum dynamics that lead to the classical behavior of the apparatus must preserve the recorded outcome. Since chaotic quantum dynamics scramble as much information as symmetries allow [22, 119–122], the states of $\mathcal{H}_{\text{ss}}$ associated with each outcome m must belong to distinct symmetry sectors. Together with the first observation, this implies that the chaotic evolution of the apparatus is well approximated by a random unitary of the form

$$U = \sum_m P_{\text{ss}}^{(m)} U_m P_{\text{ss}}^{(m)} \quad \text{for } U_m \in \text{U}(r_m), \quad (5.15)$$

where the $r_m \times r_m$ unitary U_m acts in the r_m -dimensional subspace of $\mathcal{H}_{\text{ss}}$ corresponding to the m th outcome O_m (2.3) and is drawn at random from $\text{U}(r_m)$ with uniform (Haar) measure. In the minimal Stinespring representation, each outcome has a unique state, so $r_m = 1$ and $U_m = e^{i\theta_m}$ is simply a complex phase. In the nonminimal representations relevant to experiments, U_m generically mixes states that reflect the same outcome O_m . For example, in the case of photon counting (see Sec. 3.2, every state of $\mathcal{H}_{\text{ss}}$ with n conduction electrons reflects the same outcome n , so that the Haar-random unitary U (5.15) respects a $\text{U}(1)$ symmetry corresponding to particle number.

We next show that the chaotic, symmetry-preserving unitary evolution of the apparatus—captured by Haar-averaged evolution under Eq. 5.15—leads to wavefunction collapse. Following the measurement process, the full dilated state is generically given by

$$\varrho(t) = \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho_0 P_n \otimes \sum_{j=1}^{r_m} \sum_{k=1}^{r_n} c_{m,j}^* c_{n,k} |m,j\rangle\langle n,k|_{\text{ss}}, \quad (5.16)$$

where $|m,j\rangle$ and $|n,k\rangle$ label all states corresponding to outcomes m and n , respectively, and

$\sum_{j=1}^{r_m} |c_{m,j}|^2 = 1$. Evolving under U (5.15) and ensemble averaging leads to

$$\begin{aligned}
\varrho'(t) &= \mathbb{E} \left[U \varrho(t) U^\dagger \right] \\
&= \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho_0 P_n \otimes \sum_{j,k} c_{m,j}^* c_{n,k} \sum_{\ell,\ell'=0}^{\mathcal{N}-1} P_{\text{ss}}^{(\ell)} \mathbb{E} \left[U_\ell P_\ell |m,j\rangle \langle n,k| P_{\ell'} U_{\ell'}^\dagger \right] P_{\text{ss}}^{(\ell')} \\
&= \sum_{m,n=0}^{\mathcal{N}-1} P_m \rho_0 P_n \otimes \sum_{j,k} c_{m,j}^* c_{n,k} \sum_{\ell=0}^{\mathcal{N}-1} P_{\text{ss}}^{(\ell)} \Phi_{\text{Haar},\ell}^{(1)} (P_\ell |m,j\rangle \langle n,k| P_\ell), \tag{5.17}
\end{aligned}$$

where $\Phi_{\text{Haar},\ell}^{(1)}(A) = \text{tr}(A)/r_\ell$ is the onefold Haar channel for the $r_\ell \times r_\ell$ unitary U_ℓ [118], which vanishes unless $\ell = \ell'$ and is proportional to the identity. In particular, we have that

$$\Phi_{\text{Haar},\ell}^{(1)} (P_\ell |m,j\rangle \langle n,k| P_\ell) = \frac{1}{r_\ell} \text{tr} (P_\ell |m,j\rangle \langle n,k|) = \frac{1}{r_\ell} \delta_{m,\ell} \delta_{n,\ell} \delta_{j,k}, \tag{5.18}$$

and inserting this result back into Eq. 5.17 leads to

$$\varrho'(t) = \sum_{m=0}^{\mathcal{N}-1} P_m \rho_0 P_m \otimes \sum_{j=1}^{r_m} \frac{|c_{m,j}|^2}{r_m} P_{\text{ss}}^{(m)} = \sum_{m=0}^{\mathcal{N}-1} P_m \rho_0 P_m \otimes \frac{1}{r_m} P_{\text{ss}}^{(m)}, \tag{5.19}$$

which has precisely the desired form (5.14). The final state (5.19) is a sum over outcomes m of the probability p_m (2.4) of observing outcome m times the collapsed density matrix ρ_m (5.5) times the maximally mixed state P_m/r_m on the subspace of $\mathcal{H}_{\text{ss}}$ corresponding to outcome m . Finally, in the *minimal* Stinespring representation, each subspace has dimension $r_m = 1$, so that

$$\varrho'(t) = \mathbb{E} \left[U \mathcal{U} \varrho(0) \mathcal{U}^\dagger U^\dagger \right] = \sum_{m=0}^{\mathcal{N}-1} P_m \rho_0 P_m \otimes |m\rangle \langle m|_{\text{ss}}, \tag{5.20}$$

which is a classical mixture of $\mathcal{N}$ pure quantum states corresponding to the collapsed state $\rho_m \propto P_m \rho_0 P_m$ (5.14) tensored with the projector onto the m th Stinespring state. In the context of chaos, the process (5.20) reflects relaxation to the *diagonal ensemble*, due to dephasing [113–116]. It is ubiquitous to quantum systems that thermalize (or behave classically).

While the averaging over an ensemble of statistically similar evolution operators U (5.15) on $\mathcal{H}_{\text{ss}}$ may appear unnatural, it is merely a useful—and highly successful—approximation of the underlying chaotic dynamics. Indeed, the onefold Haar channel (5.18) is *nonunitary*, though still a valid quantum channel (i.e., CPTP map). One might think that such a nonunitary channel could not possibly capture the actual unitary evolution. In practice, however, we note that the chaotic evolution of the measurement apparatus also involves interactions with an environment, which may be similarly nonunitary in its action on the system after tracing out the environment. Moreover, the ensemble-averaged evolution reproduces the decohering effects of complicated interactions with the environment, and even successfully predicts the universal behaviors of isolated, thermodynamically large systems. We also note that other random-matrix ensembles common to the literature on, e.g., quantum complexity [118] lead to the same result (5.19). For these reasons, we expect the model of decoherence in terms of the random unitary U (5.15) to provide a useful description of the chaotic dynamics associated with the classical (and thermal) nature of the apparatus [22, 113–122].

Summarizing the discussion thus far, the axioms of quantum mechanics [1–8] imply that measurements are equivalent to unitary evolution on a *dilated* Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H}_{\text{ph}} \otimes \mathcal{H}_{\text{ss}}$ (2.8), where the unitary $\mathcal{U}$ (5.9) codifies the time evolution of the system ($\mathcal{H}_{\text{ph}}$) and measurement

apparatus ($\mathcal{H}_{\text{ss}}$) during the measurement process, and the resulting density matrix (5.14) encodes all possible outcomes. We note that this formulation of measurements is entirely *deterministic*—i.e., at the level of the equations of motion, there is nothing stochastic, and all outcomes occur.

The appearance of wavefunction collapse is merely a byproduct of our classical nature—and one we engineer. For a measurement outcome to be discernible to classical beings, it must be stored in a classical thermal state—meaning a state that is stable to the underlying and chaotic quantum dynamics. The same mechanism that leads to the classical behavior of the apparatus in the first place—and hence, its utility as a detector—also guarantee that the postmeasurement state rapidly decoheres into a classical mixture of states (5.19). In the context of measurements, we associate this with wavefunction collapse—in any given experiment, we experience only one pure state $\rho_m \otimes |m\rangle\langle m|_{\text{ss}}$, sampled from ϱ_{mixed} (5.13) with probability p_m (5.3).

The primary constraint this places on the apparatus itself is that distinct outcomes correspond to different “charges” under some symmetry; in most implementations of measurements, that symmetry corresponds to particle number. Importantly, the connection to symmetries naturally explains the emergence of a “preferred basis.” This also underlies the probabilistic appearance of radioactive decay—as classical beings, we cannot observe superpositions of different numbers of particles (symmetry sectors), as they rapidly decohere into mixed states, as in Eq. 5.14.

A related point is that any “strangeness” of quantum are by design. Without assigning distinct outcomes to different symmetry charges, the state of the apparatus would not reliably indicate the outcome. Without using a classical apparatus, there would be no notion of collapse, nor an “outcome” to discern, without a second measurement involving a classical apparatus (e.g., in the thought experiment known as “Wigner’s friend”). Moreover, any nonclassical apparatus could not safely store information about the outcome, which could easily be lost to decoherence.

This understanding of wavefunction collapse appears to resolve several “paradoxes” surrounding measurements. First, there is no paradox surrounding Schrödinger’s cat: A cat is placed in a box with a radioactive isotope, whose decay triggers the release of poison; the paradox is that, until the cat is observed, its state is a superposition of alive and dead. However, a cat is classical—in this case, it acts as the detector for the spontaneous decay of the isotope (though we comment that far more ethical and efficient detectors exist). While the release of the poison creates a coherent superposition of the cat being dead (radioactive decay occurs) and alive (no decay occurs), because the cat is classical, this state rapidly decoheres into a mixed state in which the cat is alive and no decay occurred and the cat is dead and decay occurred. This collapse happens extremely rapidly as unitary dynamics increase the probability that the isotope decayed. As a result, there is no possibility of observing the cat to be in a coherent superposition of alive and dead, and thus no paradox [11, 15]. Second, there is no concern surrounding Wigner’s friend: The apparatus should be identified with the first classical system that is entangled with the quantum system of interest, since that apparatus is sufficient for collapse, and no other ingredients are required.

Finally, we note that the physical details surrounding decoherence of the apparatus beyond the toy model considered above do *not* depend directly on the measured observable, but on microscopic and nonuniversal properties of the apparatus. In contrast to the minimal Stinespring representation (5.9), there can be no general theory of, e.g., the precise time scale over which decoherence (and collapse) occurs, as this depends entirely on microscopic details that vary depending on how the measurement is implemented. However, it is likely that the time scale of decoherence is small compared to the time scale of the measurement overall, or else the apparatus would exhibit quantum effects, and would not reliably indicate the measurement outcome. Hence, we expect that decoherence-induced collapse (5.19) occurs quite rapidly, and is merely part of the process by which the outcome becomes classically available. The crucial observations are that (i) the measurement process itself is described by a *deterministic* unitary (5.9) corresponding to time evolution of the system and apparatus; (ii) the same criteria that guarantee that $\mathcal{U}$ realizes a classically readable

measurement also guarantee the appearance of wavefunction collapse via decoherence (5.14); (iii) it also explains the emergence of a preferred classical basis for the different “branches” of the universal wavefunction $|\Psi\rangle \in \mathcal{H}_{\text{dil}}$; and (iv) that collapse is unrelated to consciousness or observation.

5.4. Absence of “spooky action at a distance”

Here we discuss how the Stinespring representation (5.9) naturally resolves the Einstein-Podolsky-Rosen (EPR) “paradox” [10], in a manner consistent with results long known to the literature. Suppose that two qubits are prepared in the Bell state

$$|\text{Bell}\rangle_{a,b} = \frac{1}{\sqrt{2}} (|0\rangle_a \otimes |0\rangle_b + |1\rangle_a \otimes |1\rangle_b) , \quad (5.21)$$

and we then send qubits a and b to Alice and Bob, respectively—who are separated by some large distance r —while preserving the maximally entangled state (5.21) of the qubits [10, 123].

Suppose that Alice and Bob each measure Z on their qubits. From the measurement axiom [1–8], it would appear that whoever measures first collapses the Bell state (5.21) onto *either* $|00\rangle$ or $|11\rangle$, so that whoever measures second is guaranteed to observe the same outcome. The EPR paradox centers on the putative violation of relativistic locality—e.g., how does Bob’s qubit b “know” to collapse to the same outcome observed by Alice upon her measurement of qubit a [10]? The idea was that information about Alice’s outcome must propagate the distance r to Bob’s qubit *instantaneously*, which would violate causality. Even if one accepts that all outcomes occur in the universal (dilated) state ϱ (5.20), since the *observed* outcomes correspond to distinct classical “branches” of ϱ (i.e., Stinespring states $|m\rangle$), one might still worry that Alice’s measurement determines this classical branch and instantaneously communicates that information to Bob’s qubit.

It has been well established for decades now that there is no violation of relativistic locality (i.e., causality) in such Bell measurements, and hence, no “paradox” to resolve. This is particularly clear using the Stinespring representation, as we now show. The dilated description reveals that there is no need for instantaneous communication between entangled qubits following a measurement of one—i.e., there is no “spooky action at a distance” in the context of measuring entangled states.

We first introduce two Stinespring qubits A and B , corresponding to the measurement apparatus used by Alice and Bob, respectively. Alice’s Z_a measurement, e.g. is captured by the dilated unitary

$$\mathcal{U}_A = |0\rangle\langle 0|_a \otimes \mathbb{1}_A + |1\rangle\langle 1|_a \otimes X_A , \quad (5.22)$$

as in Eq. 2.23, and the analogous $\mathcal{U}_B$ for Bob is the same up to replacing $(a, A) \rightarrow (b, B)$. The two apparatus are initialized in the Z -basis state $|0\rangle$, so the dilated initial state is given by

$$|\Psi_i\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)_{a,b} \otimes |00\rangle_{A,B} . \quad (5.23)$$

Assuming Alice measures first, we apply $\mathcal{U}_A$ (5.22) to $|\Psi_i\rangle$ (5.23), resulting in

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle)_{a,b,A} \otimes |0\rangle_B , \quad (5.24)$$

and applying $\mathcal{U}_B$ (5.22) for Bob’s subsequent measurement leads to the final state

$$|\Psi_f\rangle = \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle)_{a,b,A,B} , \quad (5.25)$$

from which it is clear that Alice and Bob’s Z measurements are *guaranteed* to realize the same outcome, simply due to the structure of the Bell state. Because the two measurement unitaries act on disjoint Hilbert spaces, they commute, and their order is unimportant. The correlation between the measurement outcomes is a consequence of the correlations between the physical qubits; each measurement merely entangles the corresponding apparatus into the Bell state (5.21), forming a Greenberger-Horne-Zeilinger (GHZ) state [124] of the two physical qubits and two detector qubits.

Essentially, the measurements merely reveal information already encoded in the state. Regarding the postmeasurement state $|\Psi_f\rangle$ (5.25), we see that the *physical* part of the state is *unaltered* compared to the initial state $|\Psi_i\rangle$ (5.23). In fact, projective measurements never alter the physical part of the state when written the basis of the measured observable $\mathcal{O}$ (2.3); they merely entangle the state of the apparatus into the physical state in a correlated manner.

For example, if Alice measures Z_a and Bob measures X_b , the initial state is

$$|\Psi_i\rangle = \frac{1}{2} (|0, +\rangle + |0, -\rangle + |1, +\rangle - |1, -\rangle)_{a,b} \otimes |0, 0\rangle_{A,B}, \quad (5.26)$$

in the measurement basis $Z_a \otimes X_b$, where $X|\pm\rangle = \pm|\pm\rangle$. The postmeasurement state is

$$|\Psi_f\rangle = \frac{1}{2} (|0, +, 0, +\rangle + |0, -, 0, -\rangle + |1, +, 1, +\rangle - |1, -, 1, -\rangle)_{a,b,A,B}, \quad (5.27)$$

where the physical part is unaltered compared to $|\Psi_i\rangle$ (5.26). Again, the measurements commute.

Comparing Eqs. 5.26 and 5.27 also shows that measurements do *not* create “branches” of the universal wavefunction $|\Psi\rangle \in \mathcal{H}_{\text{dil}}$, but merely reveal superpositions *already present* in $|\Psi\rangle$. By writing $|\Psi\rangle$ in the basis of the observables to be measured, we see that the measurement simply entangles the state of the apparatus into $|\Psi\rangle$ without altering the latter, as in Eqs. 5.25 and 5.27. Classical information about the observed outcome emanates from the locations of *each* measurement, traveling no faster than c . The correlations between the outcomes of measurements of entangled states $|\Psi\rangle$ are guaranteed by the structure of $|\Psi\rangle$ itself, without the need for action at a distance.

Not only is there no *need* for action at a distance—“spooky” or otherwise—but we now explain why there cannot be any such action. Returning to the well separated Bell state (5.21), the expectation value of any Pauli measurement by either Alice or Bob is given by

$$\langle \sigma_{a/b}^\nu \rangle = \text{tr} \left(\sigma_{a/b}^\nu \rho_{\text{Bell}} \right) = \frac{1}{4} \text{tr} \left[\sigma_{a/b}^\nu (\mathbb{1} + X_a X_b + Z_a Z_b - Y_a Y_b) \right] = 0, \quad (5.28)$$

so that, if they repeat the experiment many times without communicating, Alice and Bob each conclude that their qubit is in the maximally mixed state $\rho_{a/b} = \mathbb{1}/2$, so that all measurement outcomes are equally likely. They can only deduce that their qubits are maximally entangled with *something else*—e.g., a thermal bath—but not that they are part of a Bell state.

In fact, without communicating, Alice and Bob cannot learn anything about the other party. If they both know that they share a Bell state, they cannot determine which of the four Bell states they share from local measurements. Even knowing which Bell state they share, they cannot determine *whether* the other party has measured yet, even knowing what they plan to measure. Alternatively, even knowing that the other party has measured, they cannot determine which operator was measured. The *only* way that Alice and Bob can see correlations between their measurements is by sending classical signals, which obey relativistic causality. The *only* way Alice and Bob can use their entangled state to transmit quantum information is again by sending a classical signal containing details of the measurements and their outcomes [21, 23]. In other words, there is no “paradox” because measuring one party in an entangled state does not transmit information or influence of any kind to the other parties—there is no action at a distance, as is now well known.

5.5. Quantum mechanics is local

We now explain that not only measurements of Bell states [10, 123], but *all* of quantum mechanics, is compatible with relativistic locality [21]. In the process, we show how the Stinespring representation of Sec. 2 implies that large swaths of quantum dynamics—including those with measurements and nonlocal feedback—obey a much stronger, *nonrelativistic* notion of locality [21] in the sense of the Lieb-Robinson theorem [111]. Such notions of locality have profound implications for quantum dynamics, phases of matter, and more [21, 111, 125].

We emphasize that we use the term “locality” exclusively to refer to the idea that objects are only influenced by events within their causal light cone. We do *not* refer to the notion of “quantum nonlocality” associated with Bell states [123], which instead refers to the fact any *classical* hidden-variables model describing quantum phenomena *must* violate relativistic locality [126, 127]. Instead, we use the term “locality” to refer to the idea that information and influence of any kind can always be traced dynamically, and never exceed the speed of light c . Moreover, in most nonrelativistic scenarios, there exists an *emergent* and stronger notion of locality.

We first review locality in the context of unitary time evolution of isolated quantum systems. Consider a lattice spin system described by a local Hamiltonian H (e.g., interactions between neighboring spins) with energy scale J and average spacing a between spins. The Lieb-Robinson bound establishes that the distance x over which quantum information can be transferred (or correlations and/or entanglement generated, etc.) in time t obeys the inequality

$$x \leq 2vt, \quad (5.29)$$

where $v \sim aJ/\hbar \ll c$ is an *emergent* speed limit on quantum information [111]. The bound comes from the fact that correlations, entanglement, and quantum information can spread out from some point no faster than v , leading to a region of size $2vt$ in time t . Formally, the bound (5.29) derives from the fact that generic commutators satisfy $||[A(x, t), B(0, 0)]|| \lesssim \exp(2vt - x)$ as $x \rightarrow \infty$.

The Lieb-Robinson bound (5.29) also applies to systems with interactions in finite local regions, exponentially decaying interactions $V(r) \sim e^{-r/\xi}$ [125], and even power-law interactions $V(r) \sim r^{-\alpha}$ for $\alpha \geq 2d + 1$ in d spatial dimensions [128–131]. A modified Lieb-Robinson bound with a nonlinear “light cone” recovers for $d < \alpha < 2d + 1$ [131]. The foregoing Hamiltonians all have an emergent notion of *nonrelativistic* locality that is even stronger (i.e., more constraining) than standard, relativistic locality. Importantly, *all* Hamiltonians H obey *relativistic* locality: When $\alpha \leq d$, $V(r)$ is mediated by a field whose excitations propagate no faster than c (e.g., the Coulomb potential).

Still, generic quantum dynamics may also involve nonunitary quantum operations, such as measurements and outcome-dependent feedback. When the outcome-dependent operations are performed in the vicinity of the measurement, the entire process is captured by a Lindblad master equation involving only *local* jump operators. Any dynamics captured by a local Lindbladian obeys the *same* Lieb-Robinson bound (5.29) as local unitary time evolution alone [112].

However, when outcome-dependent feedback is applied *nonlocally*—i.e., far from the location of measurement—the Lieb-Robinson bound (5.29) does *not* apply [21]. We allow for generic local quantum operations that are individually compatible with the usual Lieb-Robinson bound [111, 112], and further allow for the nonlocal communication of measurement outcomes; in the nonrelativistic limit where $c \rightarrow \infty$, this communication is effectively instantaneous. Using the Stinespring representation of Sec. 2, Ref. 21 established that *any* quantum task that achieves a useful task over some distance x using measurements in M regions obeys the *generalized* bound,

$$x \leq 2(M + 1)vt, \quad (5.30)$$

where t is the duration of unitary time evolution and v is the Lieb-Robinson velocity associated with that evolution (5.29). While the bound (5.30) was only proven explicitly for qubits on arbitrary

graphs and interacting with their nearest neighbor only, it is highly likely that the generalized bound is compatible with any Lieb-Robinson bound for the purely unitary part [21]. Proving this to be the case is an open direction for future work. We stress that the measurements need not be projective, and can be generalized to any quantum operations involving ancillas in $\mathcal{H}_{ss}$; however, any such operation *not* accompanied by nonlocal outcome-dependent feedback provides no enhancement [21].

We further comment that the generalized bound (5.30) does *not* derive from commutator norms, as measuring Z_i and then communicating the outcome $n \in \{0, 1\}$ and applying X_j if $n = 1$ leads to $|[X_i(0), Z_j(t)]| = 2$ even as $t \rightarrow 0$ and $x = d(i, j) \rightarrow \infty$. However, this sequence of operations does not transmit quantum information nor does it generate or meaningfully alter correlations or entanglement between i and j . The commutator is nonzero solely because classical information was transmitted. Instead, the bound (5.30) derives from comparing the reduced density matrix ρ_{ij} between sites i and j separated by x and showing that, if the bound is violated, it is arbitrarily close in trace distance to a state ρ'_{ij} that contains no entanglement or correlations between i and j . Because all useful quantum tasks inherently create entanglement and/or correlations over some distance x , the bound (5.30) constrains all useful quantum tasks [21].

Importantly, the generalized bound (5.30) implies that a finite number of measurements only leads to a *finite* enhancement to the speed v (5.29) of quantum information. This holds even in the nonrelativistic limit where $c \rightarrow \infty$ and communication of outcomes is *instantaneous*. Intuitively, the bound reflects the fact that measurements can only be used to “link up” local regions—of a resource state, teleportation protocol, or similar. The most optimal protocols involve unitary evolution for time t that generates local patches of the resource state in $M + 1$ regions. Each patch grows to maximum size $\ell(t) = 2vt$ (5.29), and the M measurements “stitch” these patches together to achieve a distance of $x \leq (M + 1)\ell(t) = 2(M + 1)vt$ (5.30). In some cases, a protocol can achieve a useful quantum task starting from a product state (e.g., $|00 \cdots 00\rangle$) in as few as $t = 2$ layers of entangling unitary gates (i.e., depth two). In the context of teleportation, the outcomes of measurements adjacent to the initial site i must be communicated to determine a feedback operation on the final site f , with $d(i, f) = x$; in the context of state preparation, measurements at the edges of the system are tied to feedback operations *throughout* the system, with maximum linear size x [21].

Essentially, there are two mechanisms for transferring quantum information (or generating correlations and/or entanglement). The first is through unitary time evolution, which usually obeys a nonrelativistic Lieb-Robinson bound (5.29) and always obeys relativistic locality [111]. The second is through the communication of measurement outcomes and accompanying feedback, which obeys the generalized bound (5.30) whenever the unitary dynamics obey the Lieb-Robinson bound (5.29) and the number of measurement regions M is finite [21]. But when $M \rightarrow \infty$ or the unitary evolution does not obey Eq. 5.29, there is no nonrelativistic notion of locality.

However, the *relativistic* notion of locality always holds. We stress that the generalized bound (5.30) holds in the *nonrelativistic* limit where $c \rightarrow \infty$; to demonstrate relativistic locality, we take c to be finite but large. First, consider an optimal protocol with $M = \mathcal{O}(x)$ measurements and unitary evolution with Lieb-Robinson velocity v (5.29) for time t_u , which we minimize. Because the protocol is optimal, all $M \rightarrow \infty$ measurements occur after all unitary evolution [21], and we assume that they occur simultaneously in the rest frame of the system. Even in this limit, at least one measurement’s outcome must be communicated a distance $x - 2vt_u$ using a signal with speed c , where the unitary dynamics shorten this distance by vt_u from both “edges” of the system. This protocol achieves a useful task over distance x in time $t \geq x/c + (1 - 2v/c)t_u$, i.e.,

$$x \leq c(t - t_u) + 2vt_u \leq 2ct, \quad (5.31)$$

meaning that information travels at speed $2v \ll c$ for time t_u and speed c for the remaining time; since $v \ll c$, this obeys relativistic locality. Even in the limit $v \rightarrow c$, we still have $x < 2ct$, so

relativistic locality is always satisfied. The reason is that the measurements themselves do not send information—that is achieved by the classical signals, which travel no faster than c .

Accordingly, *all* quantum operations obey relativistic locality. A large fraction obey a stronger, nonrelativistic notion of locality, captured by the Lieb-Robinson bound (5.29) for unitary evolution and local quantum operations, and by the generalized bound (5.30) for local quantum operations combined with classical signaling and nonlocal feedback [21, 111, 112]. Without the classical signals and the corresponding feedback, nonunitary operations provide no advantage over unitary dynamics [21, 22, 112]. The generation of correlations and/or entanglement and the transmission of quantum information involve combinations of unitary dynamics and classical signals, which individually and jointly obey relativistic locality. Hence, all quantum dynamics are local in the relativistic sense.

This not only establishes that quantum mechanics is “complete” in the sense of EPR [10], but also has extremely useful implications for quantum applications. For example, locality can be used to constrain quantum protocols, reveal useful tradeoffs (e.g., between unitary time evolution and measurements), identify optimal protocols tailored to particular hardware, and constrain the resource states compatible with a particular quantum task [21, 23], among other applications.

6. CONCLUSION

The description of quantum measurements as an entangling interaction between a physical system of interest and a measurement apparatus—as well as the unitary description (2.51) of generic quantum operations in finite dimension—have long been known to the literature [11–13, 15, 20–24, 28–33]. Nonetheless, the analytical tools most commonly used in the literature to describe measurements manifestly ignore details of the apparatus [14, 17–20], creating a disconnect between theory and experiment, leading to avoidable confusion about the nature of measurements, and complicating the description of adaptive protocols with midcircuit measurements [21–23].

In this work, we have given an overview of projective and destructive quantum measurements, their theoretical description, and their experimental implementation. In particular, we have highlighted the “Stinespring” representation of measurements in terms of a unitary $\mathcal{U}$ (2.17) acting on a dilated Hilbert space $\mathcal{H}_{\text{dil}} = \mathcal{H} \otimes \mathcal{H}_{\text{ss}}$ (2.8), and how it follows logically from all axiomatic formulations of quantum mechanics [1–8]. We have also shown how the unitary $\mathcal{U}$ relates to the standard Kraus representation (2.11) [17–20], von Neumann pointer Hamiltonian H (2.34) [32, 33], and most importantly, to *experiment*. Noting that a large number of experimental implementations culminate in one of a handful of measurement types, we consider the most prominent among these in Secs. 3 and 4. In all cases, we find that the dilated unitary $\mathcal{U}$ (5.9) is not just a bookkeeping tool but *physical*, capturing the time evolution of the system and measurement apparatus during the measurement process, in agreement with intuition from the literature.

Generally speaking, projective and destructive measurements involve forming a maximally entangled state between the physical system and a classical measurement apparatus. Consider the measurement of an observable $\mathcal{O}$ (2.3) with $\mathcal{N}$ unique eigenvalues (i.e., measurement outcomes) O_m with corresponding eigenprojectors P_m (2.1). In the *minimal* Stinespring representation, the projective measurement outputs a state that is a sum over m of the collapsed state $P_m|\psi\rangle \in \mathcal{H}$ of the system times the state $|m\rangle_{\text{ss}} \in \mathcal{H}_{\text{ss}}$ of the apparatus. Additionally, the measurement unitary $\mathcal{U}$ (5.9) is unique up to the choice of initial state $|i\rangle_{\text{ss}} \in \mathcal{H}_{\text{ss}}$ of the apparatus. In the case of *destructive* measurements—which are unitarily related to projective measurements—the final state of the physical system is not an eigenstate of the measured observable. In our consideration of physical implementations of measurements in Secs. 3 and 4, we saw that the physical unitary corresponding to the time evolution of the system and apparatus generally realizes a *nonminimal* Stinespring representation, in which $\dim(\mathcal{H}_{\text{ss}}) > \mathcal{N}$. However, upon “binning” distinct states of the apparatus

corresponding to the same observed outcome—and ignoring states that do not encode information about the outcome—one recovers the minimal Stinespring representation (5.9).

In fact, one of the most important features of the Stinespring formulation of measurements on $\mathcal{H}_{\text{dil}}$ (2.8) is its direct connection to experiment. In all of the examples of projective and destructive measurements considered in Secs. 3 and 4, we found that the Stinespring measurement unitary $\mathcal{U}$ (5.9) corresponds *exactly* to a simplified description of the actual, physical time evolution of the system and measurement apparatus during the measurement process. This direct correspondence suggests that $\mathcal{U}$ (5.9) is not only a valid bookkeeping tool, but the physical time evolution operator. This allows us to evolve operators in the Heisenberg picture through arbitrary combinations of measurements and outcome-dependent feedback [21–23]. Thus, the minimal Stinespring representation not only encodes the minimal ingredients required for the experimental implementation of a given measurement, but is a powerful analytical tool for describing adaptive protocols with midcircuit measurements.

Additionally, we have explained in Sec. 5 how to extract standard results from the Stinespring formalism, and its application to quantum protocols. We explained in Sec. 5.1 how to recover the familiar Born rule, expectation values, and other statistics from the *post*measurement state $\rho(t)$ (5.4), using Gleason’s theorem [110]. We showcased the utility of the Stinespring representation $\mathcal{U}$ (5.9) for adaptive protocols with measurements and outcome-dependent feedback in Sec. 5.2. These protocols can achieve tasks in much less time than their purely unitary counterparts [21].

In Sec. 5.3 we derived the appearance of *wavefunction collapse* via decoherence [15]. Because the state of the measurement apparatus is readable to us, it must be stable to the quantum decoherence—i.e., the chaotic, highly entangling interactions with the environment—responsible for the classical behavior of the apparatus. Moreover, the states corresponding to distinct outcomes must be *stable*; as is known from the quantum chaos literature [22, 120, 121], this requires that these states have different “charges” under some symmetry (e.g., particle number), which also provides a “classically preferred basis” [15]. This decoherence effectively leads to a mixed state (5.14)—i.e., a classical ensemble of “collapsed” pure quantum states corresponding to different outcomes. Despite being deterministic, we *experience* this process as probabilistic because we only experience one of the pure states encoded in a mixed state, according to their probabilities.

We have also explained using the Stinespring representation how *all* quantum operations—including the combination of measurements and nonlocal feedback—obey relativistic locality. In Sec. 5.4, we explain the absence of “spooky action at a distance” in measurements of a well separated Bell state [10, 21, 123]. If Alice and Bob share an entangled Bell state, there is no measurement either can perform on their own qubit to learn anything about the other’s. Alice’s measurement of her qubit does not influence or in any way transmit information to Bob’s qubit. Hence, there is simply no “action at a distance” of any kind, and “quantum nonlocality” [126, 127] is a misnomer. We also show how the act of measurement does not create new “branches” in the universal wavefunction $|\Psi\rangle \in \mathcal{H}_{\text{dil}}$. In fact, when the physical state $|\psi\rangle \in \mathcal{H}$ is written in the basis of the observable to be measured, the projective measurement process does not even alter the structure of $|\psi\rangle$, but merely entangles the state of the apparatus into the existing superposition. Moreover, in Sec. 5.5, we summarized how the Stinespring representation was used in Ref. 21 to establish a notion of locality in quantum dynamics with measurement, generalizing the Lieb-Robinson bound [111] to *generic* quantum channels. Most importantly, we also showed how this bound implies that *all quantum dynamics respect relativistic locality* (i.e., causality).

In summary, we have shown how the unitary Stinespring representation of measurements on a dilated Hilbert space that includes the detector is both conceptually transparent and practically useful. While this representation is likely intuitive to many [24], it is not commonly used in the literature, despite its utility. In addition to providing a useful connection between theory and experiment, a powerful tool for describing protocols with midcircuit measurements, and insights into more foundational questions about quantum measurements, the Stinespring formalism also paves

the way to entirely new types of results in the context of quantum computation, dynamics, and information processing. In particular, the formalism allows for the Heisenberg evolution of *operators* in the presence of arbitrary quantum operations (e.g., measurements and nonlocal feedback), which allows for the systematic study of quantum protocols, resolving questions such as what resource states can be measured to achieve a given quantum task [23], and revealing new constraints and resource tradeoffs [21]. We believe this formalism will be an invaluable tool in future studies of quantum protocols and near-term quantum technologies. In addition to forthcoming work proving the physicist’s Stinespring theorem for infinite-dimensional systems [31], we also plan to extend the Stinespring representation to generalized and weak measurements, and generic quantum operations.

ACKNOWLEDGMENTS

We thank Josh Combes, Shawn Geller, Eric Song, Daniel Spiegel, James Woodcock, and Aaron Young for useful discussions and feedback on this manuscript. DB is supported in part by NIST, the DOE Quantum Systems Accelerator (QSA) grant, ARO grant W911NF-19-1-0210, the Simons collaboration on Ultra-Quantum Matter (UQM), and NSF grants JILA-PFC PHY-2317149, QLCI-OMA-2016244, and PHY-2309135 via the Kavli Institute for Theoretical Physics (KITP); DB also acknowledges the hospitality of the KITP while parts of this work were completed. AJF is supported in part by DOE grant DE-SC0024324.

Appendix A: Details of the Stern-Gerlach experiment

Here we elaborate on technical details relevant to the Stern-Gerlach experiment discussed in Sec. 2.4.2. In App. A.1 we show that the evolution of the y coordinate can be ignored; in App. A.2, we derive the final wavefunction $\Psi_s(z, t)$ for the magnetic particle.

A.1. Time evolution of operators in the general case

Observables $\mathcal{O}$ evolve in the Heisenberg picture under a Hamiltonian H according to

$$\mathcal{O}(t) = e^{iHt/\hbar} \mathcal{O} e^{-iHt/\hbar} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{t}{i\hbar} \right)^n [\mathcal{O}, H]_n, \quad (\text{A.1})$$

where the nested commutator satisfies $[A, B]_{n+1} = [[A, B]_n, B]$ with $[A, B]_0 = A$. For the Stern-Gerlach Hamiltonian H_{SG} (2.39), the j th component of the position operator $\mathbf{x}$ evolves via

$$x_j(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[\mathcal{O}, \left(-\frac{\hbar^2}{2M} \nabla^2 - \mu_B b y \sigma^y + \mu_B (B_0 + b z) \sigma^z \right) \frac{t}{i\hbar} \right]_n = \sum_{n=0}^{\infty} \frac{1}{n!} x_j^{(n)}(t), \quad (\text{A.2})$$

where we retain the time dependence of the n th term for convenience. The first several terms are

$$x_j^{(0)}(t) = x_j \quad (\text{A.3a})$$

$$x_j^{(1)}(t) = \frac{t}{i\hbar} [x_j, H_{\text{SG}}] = \frac{it\hbar}{2M} [x_j, \partial_j^2] = \frac{t\hbar}{iM} \partial_j = \frac{t}{M} p_j \quad (\text{A.3b})$$

$$x_j^{(2)}(t) = \frac{t}{i\hbar} \left[\frac{t\hbar}{iM} \partial_j, H_{\text{SG}} \right] = \frac{\mu_B b t^2}{M} (\delta_{j,y} \sigma^y - \delta_{j,z} \sigma^z), \quad (\text{A.3c})$$

and for convenience, we now treat the evolution of y and z separately. For $y(t)$, we have

$$y^{(3)}(t) = \frac{t}{i\hbar} \left[\frac{\mu_B b t^2}{M} \sigma^y, H_{\text{SG}} \right] = \frac{\mu_B b t^3}{i\hbar M} [\sigma^y, \mu_B (B_0 + bz) \sigma^z] = \frac{2\mu_B^2 b t^3}{\hbar M} (B_0 + bz) \sigma^x \quad (\text{A.4a})$$

$$\begin{aligned} y^{(4)}(t) &= \frac{t}{i\hbar} \frac{2\mu_B^2 b t^3}{\hbar M} [(B_0 + bz) \sigma^x, H_{\text{SG}}] \\ &= \frac{2\mu_B^2 b t^4}{i\hbar^2 M} \left(-\frac{b\hbar^2}{2M} [z, \partial_z^2] \sigma^x - \mu_B b y (B_0 + bz) [\sigma^x, \sigma^y] + \mu_B (B_0 + bz)^2 [\sigma^x, \sigma^z] \right) \\ &= \frac{2\mu_B^2 b t^4}{i\hbar^2 M} \left(\frac{b\hbar^2}{M} \sigma^x \partial_z - 2i\mu_B b y (B_0 + bz) \sigma^z - 2i\mu_B (B_0 + bz)^2 \sigma^y \right) \\ &= \frac{2\mu_B^2 b^2 t^4}{\hbar M^2} p_z \sigma^x - \frac{4\mu_B^3 b t^4}{\hbar^2 M} (B_0 + bz)^2 \sigma^y - \frac{4\mu_B^3 b^2 t^4}{\hbar^2 M} (B_0 y + b y z) \sigma^z, \end{aligned} \quad (\text{A.4b})$$

where the terms above are organized by the Pauli matrices. The analogous expressions for $z(t)$ are

$$z^{(3)}(t) = \frac{t}{i\hbar} \left[-\frac{\mu_B b t^2}{M} \sigma^z, H_{\text{SG}} \right] = \frac{i\mu_B b t^3}{\hbar M} [\sigma^z, -\mu_B b y \sigma^y] = -\frac{2\mu_B^2 b^2 t^3}{\hbar M} y \sigma^x \quad (\text{A.5a})$$

$$\begin{aligned} z^{(4)}(t) &= \frac{t}{i\hbar} \left(-\frac{2\mu_B^2 b^2 t^3}{\hbar M} \right) [y \sigma^x, H_{\text{SG}}] \\ &= \frac{2i\mu_B^2 b^2 t^4}{\hbar^2 M} \left(-\frac{\hbar^2}{2M} [y, \partial_y^2] \sigma^x - \mu_B b y^2 [\sigma^x, \sigma^y] + \mu_B y (B_0 + bz) [\sigma^x, \sigma^z] \right) \\ &= \frac{2i\mu_B^2 b^2 t^4}{\hbar^2 M} \left(\frac{\hbar^2}{M} \partial_y \sigma^x - 2i\mu_B b y^2 \sigma^z - 2i\mu_B y (B_0 + bz) \sigma^y \right) \\ &= -\frac{2\mu_B^2 b^2 t^4}{\hbar M^2} p_y \sigma^x + \frac{4\mu_B^3 b^2 t^4}{\hbar^2 M} (B_0 y + b y z) \sigma^y + \frac{4\mu_B^3 b^3 t^4}{\hbar^2 M} y^2 \sigma^z, \end{aligned} \quad (\text{A.5b})$$

organized as before. Now, suppose we plan to measure z at time t , and take the initial state to be

$$|\Psi(0)\rangle = \iiint_{\mathbb{R}^3} dp_x dy dz \frac{e^{-(p_x - Mv)^2/4\delta_x^2}}{(2\pi\delta_x^2)^{1/4}} \frac{e^{-y^2/4\delta_y^2}}{(2\pi\delta_y^2)^{1/4}} \frac{e^{-(z-z_0)^2/4\delta_z^2}}{(2\pi\delta_z^2)^{1/4}} |p_x, y, z, s\rangle, \quad (\text{A.6})$$

where $\sigma^z |s\rangle = s |s\rangle$. This initial is a Gaussian wave packet centered around $y = 0$ and $z = z_0$ (in real space) and $p_x = Mv$ (in momentum space), with spin polarization s . To $\mathcal{O}(t^4)$, we have

$$\langle x \rangle(t) = vt \quad (\text{A.7a})$$

$$\langle y \rangle(t) = 0 \quad (\text{A.7b})$$

$$\langle z \rangle(t) = z_0 - \frac{\mu_B b t^2}{2M} s + \frac{\mu_B^3 b^3 t^4}{6\hbar^2 M} \delta_y^2 s, \quad (\text{A.7c})$$

meaning that the average correction to $z(t)$ due to the y dynamics only appears at $\mathcal{O}(t^4)$, and is proportional to the variance δ_y in the initial y position. By minimizing δ_y and using particles with polarized z spin s , the y dynamics can be neglected, as in Sec. 2.4.2.

A.2. Time evolution of states in the simple case

We start by applying the rightmost operator in the unitary $\mathcal{U}(t)$ (2.43) to $|\Psi(0)\rangle$ (2.42), giving

$$\sum_{s=\pm 1} c_s \int dz \varphi_z(z) |z, s\rangle \mapsto \sum_{s=\pm 1} c_s \int dz \varphi(z - st^2 \frac{b\mu_B}{M}) |z, s\rangle, \quad (\text{A.8})$$

via the same trick for exponentials of ∂_z used in the pointer-particle example (2.37) in Sec. 2.4.1, where $\varphi(z) \propto \exp(-(z - z_0)^2/4\delta^2)$ (A.6). Applying the next factor in $\mathcal{U}(t)$ (2.43) leads to

$$\begin{aligned} |\Psi(0)\rangle &\mapsto \sum_{s=\pm 1} c_s \int dz \left[\exp\left(\frac{i\hbar t}{2M} \partial_z^2\right) \right] \varphi(z - z_{s,t}) |z, s\rangle \\ &= \sum_{s=\pm 1} c_s \int dz \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\hbar t}{2iM}\right)^n \partial_z^{2n} \frac{\exp\left[-(z - z_{s,t})^2/4\delta^2\right]}{(2\pi\delta^2)^{1/4}} |z, s\rangle \\ &= \sum_{s=\pm 1} c_s \int dz \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i\hbar t}{8\delta^2 M}\right)^n H_{2n}\left(\frac{z - z_{s,t}}{2\delta}\right) \varphi(z - z_{s,t}) |z, s\rangle, \end{aligned} \quad (\text{A.9})$$

where $H_k(y)$ is the k th (physicist's) Hermite polynomial in y , and we have defined

$$z_{s,t} \equiv z_0 + \frac{b\mu_B}{M} s t^2, \quad (\text{A.10})$$

and we now simplify Eq. A.9 using a Taylor expansion of the Hermite polynomials. We define

$$\begin{aligned} F(\lambda, \tau) &= \sum_{n=0}^{\infty} \frac{1}{n!} (i\tau)^n H_{2n}(\lambda) = \sum_{n=0}^{\infty} \frac{1}{n!} (i\tau)^n \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \partial_{\lambda}^k H_{2n}(\lambda) \Big|_{\lambda=0} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (i\tau)^n \sum_{k=0}^n \frac{\lambda^k}{k!} \frac{2^k (2n)!}{(2n-k)!} H_{2n-k}(0) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (i\tau)^n \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} \frac{4^k (2n)!}{(2n-2k)!} H_{2n-2k}(0), \end{aligned} \quad (\text{A.11})$$

where we used the fact that $H_{2s+1}(0) = 0$, and we next switch the order of summation,

$$F(\lambda, \tau) = \sum_{k=0}^{\infty} \frac{(2\lambda)^{2k}}{(2k)!} \sum_{n=k}^{\infty} \frac{(i\tau)^n}{n!} \frac{(2n)!}{(2n-2k)!} H_{2(n-k)}(0), \quad (\text{A.12})$$

and the sum over n can be computed exactly to give

$$F(\lambda, \tau) = \sum_{k=0}^{\infty} \frac{(2\lambda)^{2k}}{(2k)!} \frac{(2k)!}{k!} \frac{1}{\sqrt{1+4i\tau}} \left(\frac{i\tau}{1+4i\tau}\right)^k = \frac{1}{\sqrt{1+4i\tau}} \exp\left(\frac{4i\lambda^2\tau}{1+4i\tau}\right), \quad (\text{A.13})$$

so that the initial wavefunction (2.42) evolves via $\mathcal{U}(t)$ (2.43) into

$$\Psi_s(z, t) = c_s \frac{e^{ib^2\mu_B^2 t^3/\hbar M} e^{-ist\mu_B(B_0+bz)/\hbar}}{(2\pi)^{1/4} (\delta + \frac{i\delta t}{2\delta})^{1/2}} \exp\left[-\frac{\left(z - z_0 - \frac{b\mu_B}{M} s t^2\right)^2}{4\delta^2} \frac{1}{1 + \frac{i\hbar t}{2M\delta^2}}\right], \quad (\text{2.44})$$

as claimed in the main text. This simplifies upon evaluating $p_s(z, t)$ (2.45).

Appendix B: Photon absorption

Here we discuss details of the photon-counting measurements considered in Sec. 3.2. The physical system is a bosonic mode with $\mathcal{H}_{\text{ph}} = L^2(\mathbb{R}) = \ell^2(\mathbb{N}) \cong \ell^2(\mathbb{N})$, and the apparatus consists of $\mathcal{N} \ll 1$

qubits with states $|0\rangle$ (default) and $|1\rangle$ (excited). We initialize the system in the state $|\psi\rangle$, the apparatus in the state $|\mathbf{0}\rangle = |0\rangle^{\otimes \mathcal{N}}$, and evolve under a Hamiltonian H (3.7), leading to

$$|\Psi\rangle = e^{g\tau(a\sigma_{\mathcal{N}}^+ - a^\dagger\sigma_{\mathcal{N}}^-)} \dots e^{g\tau(a\sigma_1^+ - a^\dagger\sigma_1^-)} |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}}, \quad (3.8)$$

where τ is the duration and g is the strength of the interaction between the light and the detector qubits, which are labeled according to the order in which they interact with the light.

Next, consider the interaction between the electromagnetic mode and the k th detector qubit. Assuming that the former starts in the state $|\phi\rangle$ and taking $\mathcal{N} \gg 1$ and $\tau \ll 1$, we find that

$$\begin{aligned} e^{g\tau(a\sigma_k^+ - a^\dagger\sigma_k^-)} |\phi\rangle \otimes |0\rangle_k &= e^{g\tau a \sigma_k^+} e^{-g\tau a^\dagger \sigma_k^-} e^{g^2\tau^2[a\sigma_k^+, a^\dagger\sigma_k^-]/2} (1 + \mathcal{O}(g^3\tau^3)) |\phi\rangle \otimes |0\rangle_k \\ &= e^{g\tau a \sigma_k^+} e^{-g\tau a^\dagger \sigma_k^-} e^{g^2\tau^2[(a^\dagger a + 1)|1\rangle\langle 1|_k - a^\dagger a |0\rangle\langle 0|_k]/2} |\phi\rangle \otimes |0\rangle_k + \mathcal{O}(g^3\tau^3) \\ &= e^{g\tau a \sigma_k^+} e^{-g\tau a^\dagger \sigma_k^-} e^{-g^2\tau^2 a^\dagger a/2} |\phi\rangle \otimes |0\rangle_k + \mathcal{O}(g^3\tau^3) \\ &= e^{g\tau a \sigma_k^+} e^{-g^2\tau^2 a^\dagger a/2} |\phi\rangle \otimes |0\rangle_k + \mathcal{O}(g^3\tau^3), \end{aligned} \quad (B.1)$$

so that the final state $|\Psi\rangle$ (3.8) is given by

$$|\Psi\rangle = e^{g\tau a \sigma_{\mathcal{N}}^+} e^{-g^2\tau^2 a^\dagger a/2} \dots e^{g\tau a \sigma_1^+} e^{-g^2\tau^2 a^\dagger a/2} |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} + \mathcal{O}(\mathcal{N}g^3\tau^3), \quad (B.2)$$

and we now reorder the product of operators above. We first note that

$$\exp\left(-g^2\tau^2 a^\dagger a/2\right) \exp\left(g\tau a \sigma_k^+\right) = \exp\left(e^{g^2\tau^2/2} g\tau a \sigma_k^+\right) \exp\left(-g^2\tau^2 a^\dagger a/2\right), \quad (B.3)$$

which we use to commute all $e^{-g^2\tau^2 a^\dagger a/2}$ terms to the right of all others in Eq. B.2, leading to

$$\begin{aligned} |\Psi\rangle &= e^{g\tau a \sigma_{\mathcal{N}}^+} e^{-g^2\tau^2 a^\dagger a/2} \dots e^{g\tau a \sigma_1^+} e^{-g^2\tau^2 a^\dagger a/2} |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} + \mathcal{O}(\mathcal{N}g^3\tau^3) \\ &= \left[\prod_{k=1}^{\mathcal{N}} \exp\left(e^{g^2\tau^2(\mathcal{N}-k)/2} g\tau a \otimes \sigma_k^+\right) \right] e^{-\mathcal{N}g^2\tau^2 a^\dagger a/2} |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} + \mathcal{O}(\mathcal{N}g^3\tau^3), \end{aligned} \quad (B.4)$$

where all terms in the product over qubits k commute with one another, so that we can write

$$= \exp\left(g\tau \sum_{k=1}^{\mathcal{N}} e^{g^2\tau^2(\mathcal{N}-k)/2} a \otimes \sigma_k^+\right) \exp\left(-\mathcal{N}g^2\tau^2 a^\dagger a/2\right) |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} + \mathcal{O}(\mathcal{N}g^3\tau^3), \quad (B.5)$$

which we now simplify using additional definitions. We first define the *attenuation parameter*

$$\zeta \equiv \mathcal{N}g^2\tau^2, \quad (3.9)$$

which remains finite as $\tau \rightarrow 0$ and $\mathcal{N} \rightarrow \infty$. We note that there are $\mathcal{N}$ leading corrections to Eq. B.5 with $\mathcal{O}(g^3\tau^3)$ coefficients, which may conspire to produce a correction no larger than $\mathcal{O}(\mathcal{N}g^3\tau^3) = \mathcal{O}(\zeta^{3/2}\mathcal{N}^{-1/2})$, which still vanishes as $g\tau \rightarrow 0$ and $\mathcal{N} \rightarrow \infty$. We then define

$$B^\dagger \equiv \sqrt{\frac{\zeta}{\mathcal{N}(e^\zeta - 1)}} \sum_{k=1}^{\mathcal{N}} e^{\zeta(\mathcal{N}-k)/(2\mathcal{N})} \sigma_k^+, \quad (B.6)$$

as an effective bosonic operator. The postmeasurement state $|\Psi\rangle$ (B.5) now becomes

$$|\Psi\rangle = \exp\left[\left(e^\zeta - 1\right)^{1/2} a \otimes B^\dagger\right] \exp\left[-\frac{\zeta}{2} a^\dagger a\right] |\psi\rangle_{\text{ph}} \otimes |\mathbf{0}\rangle_{\text{ss}} + \mathcal{O}(\zeta^{3/2}\mathcal{N}^{-1/2}), \quad (B.7)$$

which reproduces Eq. 3.10 upon reordering the two exponential terms above.

The lowering operator B (B.6) only realizes a bosonic annihilation operator in the limit $\mathcal{N} \rightarrow \infty$ with ζ (3.9) finite. The standard commutation relation for arbitrary $\mathcal{N}$ and τ is given by

$$\begin{aligned} [B, B^\dagger] &= \frac{g^2 \tau^2}{e^{\mathcal{N} g^2 \tau^2} - 1} \sum_{k, k'=1}^{\mathcal{N}} e^{g^2 \tau^2 (2\mathcal{N} - k - k')/2} [\sigma_k^-, \sigma_{k'}^+] \\ &= -\frac{g^2 \tau^2}{e^{\mathcal{N} g^2 \tau^2} - 1} \sum_{k=1}^{\mathcal{N}} e^{g^2 \tau^2 (\mathcal{N} - k)} Z_k, \end{aligned} \quad (\text{B.8})$$

and we rewrite $Z_k = 2n_k - 1$, where $n_k = |1\rangle\langle 1|_k$. As discussed in Secs. 2.3.5 and 3.2, all states with m qubits in the state $|1\rangle$ reflect the same outcome, so we assume the qubits are excited in order of ascending k without loss of generality. The above then becomes

$$\begin{aligned} [B, B^\dagger] &= \frac{g^2 \tau^2}{e^{\mathcal{N} g^2 \tau^2} - 1} \sum_{k=1}^{\mathcal{N}} e^{g^2 \tau^2 (\mathcal{N} - k)} (1 - 2n_k) \\ &= \frac{g^2 \tau^2}{e^{\mathcal{N} g^2 \tau^2} - 1} \sum_{n=0}^{\mathcal{N}} \left(\sum_{k=1}^{\mathcal{N}} e^{g^2 \tau^2 (\mathcal{N} - k)} - 2 \sum_{k=1}^n e^{g^2 \tau^2 (\mathcal{N} - k)} \right) |n\rangle\langle n|_{\text{ss}} \\ &= \frac{g^2 \tau^2}{e^{g^2 \tau^2} - 1} \sum_{n=0}^{\mathcal{N}} \left(1 - 2 \frac{1 - e^{-ng^2 \tau^2}}{1 - e^{-\mathcal{N} g^2 \tau^2}} \right) |n\rangle\langle n|_{\text{ss}}, \end{aligned} \quad (\text{B.9})$$

and taking the limits $\tau \rightarrow 0$ and $\mathcal{N} \rightarrow \infty$ in either order while keeping ζ (3.9) fixed gives

$$[B, B^\dagger] = \sum_{n=0}^{\infty} |n\rangle\langle n|_{\text{ss}} = \mathbf{1}_{\text{ss}}, \quad (\text{B.10})$$

so that B is a bosonic annihilation operator in this limit. If we only take $\mathcal{N} \rightarrow \infty$, the above commutator is instead $g^2 \tau^2 (e^{g^2 \tau^2} - 1)^{-1} \mathbf{1}$; if we only take $\tau \rightarrow 0$, the commutator is instead $\sum_{n=0}^{\mathcal{N}} (1 - 2n/\mathcal{N}) |n\rangle\langle n|$. Hence, both limits are needed, which means that ζ (3.9) must be *finite*.

Appendix C: Matrix element for homodyne detection

Here we calculate the matrix element 3.20 central to the discussion in Sec. 3.3. This calculation also appears in the literature (see, e.g., Ref. 52). The matrix element of interest is

$$f(N, D) = \langle N, D | U_{\text{BS}} | \Psi_0 \rangle \quad (\text{C.1})$$

where $|\Psi_0\rangle = |\psi\rangle_A |\beta\rangle_B$, with $|\psi\rangle$ arbitrary and $|\beta\rangle$ (3.15) a coherent state. The Fock states $|N, D\rangle$ correspond to the sum N and difference D of the occupation numbers of the A and B modes, i.e.,

$$|N, D\rangle = \frac{1}{\sqrt{(N/2 + D)!}} \frac{1}{\sqrt{(N/2 - D)!}} (a^\dagger)^{N/2+D} (b^\dagger)^{N/2-D} |00\rangle, \quad (\text{3.19})$$

and the unitary U_{BS} describing time evolution within the beam splitter (see Fig. 2) is given by

$$U_{\text{BS}} = \exp\left(\frac{\pi}{4}(a^\dagger b - ab^\dagger)\right), \quad (\text{3.16})$$

so that the following properties hold for the annihilation operators for the two modes:

$$a U_{\text{BS}} = U_{\text{BS}} \frac{(a+b)}{\sqrt{2}} \quad (\text{C.2a})$$

$$b U_{\text{BS}} = U_{\text{BS}} \frac{(b-a)}{\sqrt{2}}, \quad (\text{C.2b})$$

and using the fact that $b|\beta\rangle_B = \beta|\beta\rangle_B$, the matrix element $f(N, D)$ (C.1) takes the exact form

$$f(N, D) = \frac{e^{-|\beta|^2/2} \langle 0 | (\beta + a)^{N/2+D} (\beta - a)^{N/2-D} |\psi\rangle_A}{\sqrt{2^N (N/2 + D)! (N/2 - D)!}}, \quad (\text{C.3})$$

using Eq. 3.15 for $\langle 0|\beta\rangle_B$, the relation $\langle 0|_A \langle 0|_B U_{\text{BS}} = \langle 0|_A \langle 0|_B$, and other relations above.

Assuming that the B-mode occupation $n_B \sim |\beta|^2 \gg 1$ is large (the precise value must be determined self-consistently and depends on n_A), we can rewrite $f(N, D)$ (C.3) as

$$\begin{aligned} &= \langle 0 | (\beta + a)^{N/2+D} (\beta - a)^{N/2-D} |\psi\rangle_A \\ &= |\beta|^N e^{-i\phi N} \langle 0 | \exp \left[\left(\frac{N}{2} + D \right) \log \left(1 + \frac{a e^{i\phi}}{|\beta|} \right) + \left(\frac{N}{2} - D \right) \log \left(1 - \frac{a e^{i\phi}}{|\beta|} \right) \right] |\psi\rangle_A \\ &\approx |\beta|^N e^{-i\phi N} \langle 0 | \exp \left(\frac{2D}{|\beta|} e^{i\phi} a - \frac{N}{2|\beta|^2} (e^{i\phi} a)^2 + \dots \right) |\psi\rangle_A, \end{aligned} \quad (\text{C.4})$$

where we have ignored higher-order terms in the arguments of the logarithms and used the definition $\beta = |\beta| e^{-i\phi}$. In the limit identified above, the matrix element $f(N, D)$ (C.3) becomes

$$f(N, D) = \frac{|\beta|^N e^{-|\beta|^2/2} e^{-i\phi N}}{\sqrt{2^N (N/2 + D)! (N/2 - D)!}} \langle 0 | \exp \left(\frac{2D}{\beta} a - \frac{N}{2\beta^2} a^2 \right) |\psi\rangle_A, \quad (\text{C.5})$$

which, as a function of N , is peaked around $N \sim |\beta|^2$. Expanding about this value leads to

$$f(N, D) = \frac{e^{i\phi N}}{|\beta| \sqrt{\pi}} e^{-D^2/|\beta|^2} e^{(N-|\beta|^2)^2/4N} \underbrace{\langle 0 | \exp \left(-\frac{1}{2} (e^{i\phi} a)^2 + \frac{2D}{|\beta|} (e^{i\phi} a) \right) |\psi\rangle}_{\langle P_\phi |}, \quad (\text{C.6})$$

and we now consider the properties of the state $|P_\phi\rangle$ above, defined by

$$|P_\phi\rangle = \exp \left[-\frac{1}{2} (e^{-i\phi} a^\dagger)^2 + \frac{2D}{|\beta|} (e^{-i\phi} a^\dagger) \right] |0\rangle \equiv \mathcal{G} |0\rangle, \quad (\text{C.7})$$

where, for notational convenience, we have introduced the operator

$$\mathcal{G} \equiv \exp \left[-\frac{1}{2} (e^{-i\phi} a^\dagger)^2 + \frac{2D}{|\beta|} (e^{-i\phi} a^\dagger) \right]. \quad (\text{C.8})$$

We first note that $|P_\phi\rangle$ is an eigenstate of the quadrature operator x_ϕ (3.14), since

$$\mathcal{G}^{-1} a^\dagger \mathcal{G} = a^\dagger \quad \text{and} \quad \mathcal{G}^{-1} a \mathcal{G} = a - a^\dagger e^{-2i\phi} + \frac{2D}{|\beta|} e^{-i\phi}, \quad (\text{C.9})$$

so that applying x_ϕ (3.14) to $|P_\phi\rangle$ (C.7) leads to

$$x_\phi |P_\phi\rangle = \left(\frac{e^{i\phi} a + e^{-i\phi} a^\dagger}{\sqrt{2}} \right) \mathcal{G} |0\rangle = \mathcal{G} \left(\frac{a e^{-i\phi}}{\sqrt{2}} + \frac{\sqrt{2}D}{|\beta|} \right) |0\rangle = \left(\frac{D\sqrt{2}}{|\beta|} \right) |P_\phi\rangle, \quad (\text{C.10})$$

meaning that $|P_\phi\rangle$ (C.7) is an eigenstate of x_ϕ (3.14) with eigenvalue $x_\phi(D) = 2D/|\beta|$. Thus, $|P_\phi\rangle$ is proportional to an eigenstate $|x_\phi(D)\rangle = \mathcal{N}^{-1}(D) |P_\phi\rangle$ that is part of an orthonormal set, i.e.,

$$\langle x_\phi(D)|x_\phi(D')\rangle = \delta(x_\phi(D) - x_\phi(D')) \propto \delta(D - D'). \quad (\text{C.11})$$

We then obtain the “normalization factor” $\mathcal{N}$ by projecting $|P_\phi\rangle = \mathcal{N}|x_\phi(D)\rangle$ onto $|0\rangle$,

$$\langle 0|P_\phi\rangle \stackrel{!}{=} 1 = \mathcal{N} \langle 0|x_\phi(D)\rangle = \frac{\mathcal{N}}{\pi^{1/4}} \exp\left[-\frac{x_\phi^2(D)}{2}\right] = \frac{\mathcal{N}}{\pi^{1/4}} e^{-D^2/|\beta|^2}, \quad (\text{C.12})$$

which we then insert into Eq. C.6 to recover the result claimed in the main text,

$$\begin{aligned} f(N, D) &= \frac{e^{i\phi N}}{|\beta| \pi^{1/4}} \exp\left[\frac{(N - |\beta|^2)^2}{4N}\right] \langle x_\phi^D|\psi\rangle \\ &= \frac{e^{i\phi N}}{|\beta| \pi^{1/4}} \exp\left[\frac{(N - |\beta|^2)^2}{4N}\right] \psi_\phi\left(\frac{D\sqrt{2}}{|\beta|}\right). \end{aligned} \quad (\text{3.20})$$

To determine the conditions under which the foregoing derivation is valid, we consider the first two terms that we neglected in the argument of the exponential in Eq. C.4, which are proportional to $D|\beta|^{-3}(a)^3$ and $N|\beta|^{-4}(a)^4$, respectively. However, we note that $N \approx |\beta|^2$ and, in typical states, $a \sim \sqrt{\langle a^\dagger a \rangle}$ (up to a phase, where the average is taken with respect to the initial physical state $|\psi\rangle_A$). Hence, we can safely neglect these higher-order terms when $|\beta| \gg \langle \psi|a^\dagger a|\psi\rangle$ and $D \ll |\beta|^3/\langle a^\dagger a \rangle^{3/2}$. Since $x_\phi(D) \propto D/|\beta|$ extends up to a distance in phase space of size $\sqrt{\langle a^\dagger a \rangle}$, the latter condition also implies that $|\beta| \gg \langle \psi|a^\dagger a|\psi\rangle$, in agreement with Ref. 66.

-
- [1] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, U.K., 1930).
 - [2] G. Birkhoff and J. V. Neumann, The logic of quantum mechanics, *Ann. Math.* **37**, 823 (1936).
 - [3] L. Hardy, Quantum theory from five reasonable axioms (2001), [arXiv:quant-ph/0101012 \[quant-ph\]](#).
 - [4] C. A. Fuchs, Quantum mechanics as quantum information (and only a little more) (2002), [arXiv:quant-ph/0205039 \[quant-ph\]](#).
 - [5] G. Mackey, *Mathematical Foundations of Quantum Mechanics* (Dover, 2004).
 - [6] A. Wilce, Four and a half axioms for finite-dimensional quantum mechanics (2009), [arXiv:0912.5530 \[quant-ph\]](#).
 - [7] L. Masanes and M. P. Müller, A derivation of quantum theory from physical requirements, *New J. Phys.* **13**, 063001 (2011).
 - [8] A. Kapustin, Is quantum mechanics exact?, *J. Math. Phys.* **54**, 062107 (2013).
 - [9] A. Einstein, Letter to Max Born, 1926, in *The Born-Einstein Letters* (Walker and Company, New York, 1971) translated by Irene Born.
 - [10] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, *Phys. Rev.* **47**, 777 (1935).
 - [11] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement*, Vol. 53 (Princeton University Press, US, 2014).
 - [12] H. Margenau, Measurements in quantum mechanics, *Ann. Phys.* **23**, 469 (1963).
 - [13] A. Peres, When is a quantum measurement?, *Amer. J. Phys.* **54**, 688 (1986).
 - [14] G. Lindblad, Entropy, information and quantum measurements, *Commun. Math. Phys.* **33**, 305 (1973).
 - [15] M. Schlosshauer, *Decoherence and the Quantum-To-Classical Transition*, Frontiers Collection (Springer Berlin, Heidelberg, Germany, 2007).
 - [16] C. A. Brasil, F. F. Fanchini, and R. d. J. Napolitano, A simple derivation of the Lindblad equation, *Revista Brasileira de Ensino de Física* **35**, 01 (2013).
 - [17] K.-E. Hellwig and K. Kraus, Pure operations and measurements, *Commun. Math. Phys.* **11**, 214 (1969).

- [18] K. Kraus, General state changes in quantum theory, *Ann. Phys.* **64**, 311 (1971).
- [19] K. Kraus, Measuring processes in quantum mechanics I. Continuous observation and the watchdog effect, *Found. Phys.* **11**, 547 (1981).
- [20] K. Kraus, *States, Effects, and Operations. Fundamental Notions of Quantum Theory.*, Lecture Notes in Physics (Springer Berlin, Heidelberg, Germany, 1983).
- [21] A. J. Friedman, C. Yin, Y. Hong, and A. Lucas, Locality and error correction in quantum dynamics with measurement, arXiv <https://doi.org/10.48550/arXiv.2206.09929> (2022), arXiv:2206.09929 [quant-ph].
- [22] A. J. Friedman, O. Hart, and R. Nandkishore, Measurement-induced phases of matter require feedback (2022), arXiv:2210.07256 [quant-ph].
- [23] Y. Hong, D. T. Stephen, and A. J. Friedman, Quantum teleportation implies symmetry-protected topological order (2023), arXiv:2310.12227 [quant-ph].
- [24] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).
- [25] R. Jozsa, An introduction to measurement-based quantum computation (2005), arXiv:quant-ph/0508124 [quant-ph].
- [26] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, UK, 2003).
- [27] M. Takesaki, *Theory of Operator Algebras I* (Springer New York, New York, US, 2012).
- [28] W. F. Stinespring, Positive functions on C^* -algebras, *Proc. Amer. Math. Soc.* **6**, 211 (1955).
- [29] M.-D. Choi, Completely positive linear maps on complex matrices, *Lin. Alg. Appl.* **10**, 285 (1975).
- [30] M. M. Wolf, *Quantum Channels and Operations — Guided Tour* (2012), graue Literatur.
- [31] A. J. Friedman and J. Woodcock, Dilation theorems for quantum operations (2024), to appear.
- [32] J. von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition*, edited by R. T. Beyer and N. A. Wheeler (Princeton University Press, US, 2018).
- [33] J. Preskill, The Physics of Quantum Information (2022), arXiv:2208.08064 [quant-ph].
- [34] K. R. Davidson, C^* -Algebras by Example, in *Fields Institute Monographs*, Vol. 6 (AMS, 1996).
- [35] L. G. Brown, Extensions of AF Algebras: The Projection Lifting Problem, in *Operator Algebras and Applications, Part 1*, Proceedings of Symposia in Pure Mathematics, Vol. 38 (AMS, 1982).
- [36] O. Bratteli, Inductive limits of finite-dimensional C^* -algebras, *Trans. Amer. Math. Soc.* **171**, 195 (1972).
- [37] P. D. Drummond and Z. Ficek, *Quantum Squeezing* (Springer Berlin, Heidelberg, Germany, 2004).
- [38] K. Jacobs, *Quantum Measurement Theory and its Applications* (Cambridge University Press, 2014).
- [39] W. Gerlach and O. Stern, Der experimentelle Nachweis des magnetischen Moments des Silberatoms, *Zeitschrift für Physik* **8**, 110 (1922).
- [40] W. Gerlach and O. Stern, Das magnetische Moment des Silberatoms, *Zeitschrift für Physik* **9**, 353 (1922).
- [41] W. Gerlach and O. Stern, Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld, *Zeitschrift für Physik* **9**, 349 (1922).
- [42] P. Busch, Translation of “Die Messung quantenmechanischer Operatoren” by E.P. Wigner (2010), arXiv:1012.4372 [quant-ph].
- [43] H. Araki and M. M. Yanase, Measurement of quantum mechanical operators, *Phys. Rev.* **120**, 622 (1960).
- [44] M. Le Bellac, *Quantum Physics*, edited by P. Forcrand-Millard (Cambridge University Press, 2006).
- [45] P. Ehrenfest, Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik, *Zeitschrift für Physik* **45**, 455 (1927).
- [46] W. Magnus, On the exponential solution of differential equations for a linear operator, *CPAM* **7**, 649 (1954).
- [47] G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J. M. Raimond, and S. Haroche, Seeing a single photon without destroying it, *Nature* **400**, 239 (1999).
- [48] A. Reiserer, S. Ritter, and G. Rempe, Nondestructive detection of an optical photon, *Science* **342**, 1349 (2013).
- [49] D. Niemietz, P. Farrera, S. Langenfeld, and G. Rempe, Nondestructive detection of photonic qubits, *Nature* **591**, 570 (2021).
- [50] R. J. Glauber, The quantum theory of optical coherence, *Phys. Rev.* **130**, 2529 (1963).
- [51] S. Olivares, A. Allevi, G. Caiazzo, M. G. A. Paris, and M. Bondani, Quantum tomography of light

- states by photon-number-resolving detectors, *New J. Phys.* **21**, 103045 (2019).
- [52] J. Combes and A. P. Lund, Homodyne measurement with a Schrödinger cat state as a local oscillator, *Phys. Rev. A* **106**, 063706 (2022).
 - [53] H. P. Yuen and J. H. Shapiro, Quantum statistics of homodyne and heterodyne detection, in *Coherence and Quantum Optics IV*, edited by L. Mandel and E. Wolf (Springer, Boston, US, 1978).
 - [54] M. J. Collett, R. Loudon, and C. Gardiner, Quantum theory of optical homodyne and heterodyne detection, *J. Mod. Opt.* **34**, 881 (1987).
 - [55] A. Barchielli, Direct and heterodyne detection and other applications of quantum stochastic calculus to quantum optics, *J. Eur. Opt. Soc. B Quant. Opt.* **2**, 423 (1990).
 - [56] H. M. Wiseman and G. J. Milburn, Quantum theory of field-quadrature measurements, *Phys. Rev. A* **47**, 642 (1993).
 - [57] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, *Rev. Mod. Phys.* **81**, 299 (2009).
 - [58] P. N. J. Dennis, *Photodetectors: An Introduction to Current Technology* (Springer US, 1986).
 - [59] B. R. Mollow, Quantum theory of field attenuation, *Phys. Rev.* **168**, 1896 (1968).
 - [60] M. O. Scully and W. E. Lamb, Quantum theory of an optical maser. III. Theory of photoelectron counting statistics, *Phys. Rev.* **179**, 368 (1969).
 - [61] M. Srinivas and E. Davies, Photon counting probabilities in quantum optics, *Optica Acta: Int. J. Opt.* **28**, 981 (1981).
 - [62] W. Chmura, A quantum open-systems theory approach to photodetection, *J. Mod. Opt.* **34**, 455 (1987).
 - [63] M. Fleischhauer and D. G. Welsch, Nonperturbative approach to multimode photodetection, *Phys. Rev. A* **44**, 747 (1991).
 - [64] C. W. Gardiner and P. Zoller, *Quantum Noise*, 3rd ed., Springer Series in Synergetics (Springer, 2004).
 - [65] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University Press, UK, 2020).
 - [66] T. Tyc and B. C. Sanders, Operational formulation of homodyne detection, *J. Phys. A Math. Gen.* **37**, 7341 (2004).
 - [67] B. Yurke, S. L. McCall, and J. R. Klauder, SU(2) and SU(1,1) interferometers, *Phys. Rev. A* **33**, 4033 (1986).
 - [68] S. Prasad, M. O. Scully, and W. Martienssen, A quantum description of the beam splitter, *Opt. Commun.* **62**, 139 (1987).
 - [69] S. Wallentowitz and W. Vogel, Unbalanced homodyning for quantum state measurements, *Phys. Rev. A* **53**, 4528 (1996).
 - [70] A. Cives-Esclop, A. Luis, and L. Sánchez-Soto, Unbalanced homodyne detection with a weak local oscillator, *Optics Communications* **175**, 153 (2000).
 - [71] B. Kühn and W. Vogel, Unbalanced homodyne correlation measurements, *Phys. Rev. Lett.* **116**, 163603 (2016).
 - [72] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum, *Phys. Rev. Lett.* **70**, 1244 (1993).
 - [73] J. H. Shapiro, The quantum theory of optical communications, *IEEE J. Quant. Electr.* **15**, 1547 (2009).
 - [74] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, Quantum dynamics of single trapped ions, *Rev. Mod. Phys.* **75**, 281 (2003).
 - [75] W. S. Bakr, A. Peng, M. E. Tai, R. Ma, J. Simon, J. I. Gillen, S. Fölling, L. Pollet, and M. Greiner, Probing the superfluid-to-Mott insulator transition at the single-atom level, *Science* **329**, 547 (2010).
 - [76] J. F. Sherson, C. Weitenberg, M. Endres, M. Cheneau, I. Bloch, and S. Kuhr, Single-atom-resolved fluorescence imaging of an atomic Mott insulator, *Nature* **467**, 68 (2010).
 - [77] I. Bloch, J. Dalibard, and S. Nascimbène, Quantum simulations with ultracold quantum gases, *Nature Phys.* **8**, 267 (2012).
 - [78] H. Ott, Single-atom detection in ultracold quantum gases: A review of current progress, *Rep. Prog. Phys.* **79**, 054401 (2016).
 - [79] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, The nitrogen-vacancy colour centre in diamond, *Physics Reports* **528**, 1 (2013).
 - [80] J. F. Barry, J. M. Schloss, E. Bauch, M. J. Turner, C. A. Hart, L. M. Pham, and R. L. Walsworth, Sensitivity optimization for NV-diamond magnetometry, *Rev. Mod. Phys.* **92**, 015004 (2020).
 - [81] W. Nagourney, J. Sandberg, and H. Dehmelt, Shelved optical electron amplifier: Observation of

- quantum jumps, *Phys. Rev. Lett.* **56**, 2797 (1986).
- [82] A. H. Myerson, D. J. Szwer, S. C. Webster, D. T. C. Allcock, M. J. Curtis, G. Imreh, J. A. Sherman, D. N. Stacey, A. M. Steane, and D. M. Lucas, High-fidelity readout of trapped-ion qubits, *Phys. Rev. Lett.* **100**, 200502 (2008).
 - [83] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, Quantum spin dynamics and entanglement generation with hundreds of trapped ions, *Science* **352**, 1297 (2016).
 - [84] L. E. de Clercq, H.-Y. Lo, M. Marinelli, D. Nadlinger, R. Oswald, V. Negnevitsky, D. Kienzler, B. Keitch, and J. P. Home, Parallel transport quantum logic gates with trapped ions, *Phys. Rev. Lett.* **116**, 080502 (2016).
 - [85] K. Cui, J. Valencia, K. T. Boyce, E. R. Clements, D. R. Leibbrandt, and D. B. Hume, Scalable quantum logic spectroscopy, *Phys. Rev. Lett.* **129**, 193603 (2022).
 - [86] M. Kwon, M. F. Ebert, T. G. Walker, and M. Saffman, Parallel low-loss measurement of multiple atomic qubits, *Phys. Rev. Lett.* **119**, 180504 (2017).
 - [87] M. Martinez-Dorantes, W. Alt, J. Gallego, S. Ghosh, L. Ratschbacher, Y. Völzke, and D. Meschede, Fast nondestructive parallel readout of neutral atom registers in optical potentials, *Phys. Rev. Lett.* **119**, 180503 (2017).
 - [88] J. A. Hines, S. V. Rajagopal, G. L. Moreau, M. D. Wahrman, N. A. Lewis, O. Marković, and M. Schleier-Smith, Spin squeezing by Rydberg dressing in an array of atomic ensembles, *Phys. Rev. Lett.* **131**, 063401 (2023).
 - [89] M. Martinez-Dorantes, W. Alt, J. Gallego, S. Ghosh, L. Ratschbacher, and D. Meschede, State-dependent fluorescence of neutral atoms in optical potentials, *Phys. Rev. A* **97**, 023410 (2018).
 - [90] J. Bochmann, M. Mücke, C. Guhl, S. Ritter, G. Rempe, and D. L. Moehring, Lossless state detection of single neutral atoms, *Phys. Rev. Lett.* **104**, 203601 (2010).
 - [91] M. L. Terraciano, R. Olson Knell, D. G. Norris, J. Jing, A. Fernández, and L. A. Orozco, Photon burst detection of single atoms in an optical cavity, *Nature Physics* **5**, 480 (2009).
 - [92] W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, A quantum-gas microscope for detecting single atoms in a Hubbard-regime optical lattice, *Nature* **462**, 74 (2009).
 - [93] M. Endres, M. Cheneau, T. Fukuhara, C. Weitenberg, P. Schauß, C. Gross, L. Mazza, M. C. Bañuls, L. Pollet, I. Bloch, and S. Kuhr, Single-site- and single-atom-resolved measurement of correlation functions, *Appl. Phys. B* **113**, 27 (2013).
 - [94] N. Schlosser, G. Reymond, I. Protsenko, and P. Grangier, Sub-Poissonian loading of single atoms in a microscopic dipole trap, *Nature* **411**, 1024 (2001).
 - [95] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, H. Levine, G. Semeghini, M. Greiner, V. Vuletić, and M. D. Lukin, High-fidelity parallel entangling gates on a neutral-atom quantum computer, *Nature* **622**, 268 (2023).
 - [96] A. W. Young, W. J. Eckner, W. R. Milner, D. Kedar, M. A. Norcia, E. Oelker, N. Schine, J. Ye, and A. M. Kaufman, Half-minute-scale atomic coherence and high relative stability in a tweezer clock, *Nature* **588**, 408 (2020).
 - [97] D. Barredo, V. Lienhard, S. de Léséleuc, T. Lahaye, and A. Browaeys, Synthetic three-dimensional atomic structures assembled atom by atom, *Nature* **561**, 79 (2018).
 - [98] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, *Rev. Mod. Phys.* **93**, 025005 (2021).
 - [99] E. Jaynes and F. Cummings, Comparison of quantum and semiclassical radiation theories with application to the beam maser, *Proceedings of the IEEE* **51**, 89 (1963).
 - [100] T. Puppe, I. Schuster, A. Grothe, A. Kubanek, K. Murr, P. W. H. Pinkse, and G. Rempe, Trapping and observing single atoms in a blue-detuned intracavity dipole trap, *Phys. Rev. Lett.* **99**, 013002 (2007).
 - [101] M. H. Schleier-Smith, I. D. Leroux, and V. Vuletić, States of an ensemble of two-level atoms with reduced quantum uncertainty, *Phys. Rev. Lett.* **104**, 073604 (2010).
 - [102] K. C. Cox, G. P. Greve, J. M. Weiner, and J. K. Thompson, Deterministic squeezed states with collective measurements and feedback, *Phys. Rev. Lett.* **116**, 093602 (2016).
 - [103] O. Hosten, N. J. Engelsen, R. Krishnakumar, and M. A. Kasevich, Measurement noise 100 times lower than the quantum-projection limit using entangled atoms, *Nature* **529**, 505 (2016).
 - [104] J. M. Robinson, M. Miklos, Y. M. Tso, C. J. Kennedy, T. Bothwell, D. Kedar, J. K. Thompson, and

- J. Ye, Direct comparison of two spin-squeezed optical clocks below the quantum projection noise limit (2022), [arXiv:2211.08621 \[quant-ph\]](#).
- [105] J. Volz, R. Gehr, G. Dubois, J. Estève, and J. Reichel, Measurement of the internal state of a single atom without energy exchange, *Nature* **475**, 210 (2011).
 - [106] P. J. Windpassinger, D. Oblak, P. G. Petrov, M. Kubasik, M. Saffman, C. L. G. Alzar, J. Appel, J. H. Müller, N. Kjærgaard, and E. S. Polzik, Nondestructive probing of Rabi oscillations on the cesium clock transition near the standard quantum limit, *Phys. Rev. Lett.* **100**, 103601 (2008).
 - [107] M. R. Andrews, M.-O. Mewes, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, Direct, nondestructive observation of a Bose condensate, *Science* **273**, 84 (1996).
 - [108] C. C. Bradley, C. A. Sackett, and R. G. Hulet, Bose-Einstein condensation of lithium: Observation of limited condensate number, *Phys. Rev. Lett.* **78**, 985 (1997).
 - [109] R. Meppelink, R. A. Rozendaal, S. B. Koller, J. M. Vogels, and P. van der Straten, Thermodynamics of Bose-Einstein-condensed clouds using phase-contrast imaging, *Phys. Rev. A* **81**, 053632 (2010).
 - [110] A. Gleason, Measures on the closed subspaces of a Hilbert space, *Indiana Univ. Math. J.* **6**, 885 (1957).
 - [111] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, *Commun. Math. Phys.* **28**, 251 (1972).
 - [112] D. Poulin, Lieb-Robinson bound and locality for general Markovian quantum dynamics, *Phys. Rev. Lett.* **104**, 190401 (2010).
 - [113] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
 - [114] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
 - [115] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, *Nature* **452**, 854 (2008).
 - [116] B. Doyon, Thermalization and pseudolocality in extended quantum systems, *Commun. Math. Phys.* **351**, 155 (2017).
 - [117] O. Bohigas, M. J. Giannoni, and C. Schmit, Characterization of chaotic quantum spectra and universality of level fluctuation laws, *Phys. Rev. Lett.* **52**, 1 (1984).
 - [118] D. A. Roberts and B. Yoshida, Chaos and complexity by design, *JHEP* **2017** (4).
 - [119] A. Nahum, S. Vijay, and J. Haah, Operator spreading in random unitary circuits (2017), [arXiv:1705.08975 \[cond-mat.str-el\]](#).
 - [120] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation (2017), [arXiv:1710.09827 \[cond-mat.stat-mech\]](#).
 - [121] A. J. Friedman, A. Chan, A. De Luca, and J. T. Chalker, Spectral statistics and many-body quantum chaos with conserved charge, *Phys. Rev. Lett.* **123**, 210603 (2019).
 - [122] H. Singh, B. A. Ware, R. Vasseur, and A. J. Friedman, Subdiffusion and many-body quantum chaos with kinetic constraints, *Phys. Rev. Lett.* **127**, 230602 (2021).
 - [123] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics Physique Fizika* **1**, 195 (1964).
 - [124] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell's theorem, in *Bell's theorem, quantum theory and conceptions of the universe* (Springer, 1989) pp. 69–72.
 - [125] M. B. Hastings, Locality in quantum systems (2010), [arXiv:1008.5137 \[math-ph\]](#).
 - [126] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, *Found. Phys.* **24**, 379 (1994).
 - [127] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 419 (2014).
 - [128] M. Foss-Feig, Z.-X. Gong, C. W. Clark, and A. V. Gorshkov, Nearly linear light cones in long-range interacting quantum systems, *Phys. Rev. Lett.* **114**, 157201 (2015).
 - [129] C.-F. Chen and A. Lucas, Finite speed of quantum scrambling with long-range interactions, *Phys. Rev. Lett.* **123**, 250605 (2019).
 - [130] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas, Hierarchy of linear light cones with long-range interactions (2020), [arXiv:2001.11509 \[quant-ph\]](#).
 - [131] C.-F. Chen, A. Lucas, and C. Yin, Speed limits and locality in many-body quantum dynamics, *Rep. Prog. Phys.* **86**, 116001 (2023).