Asymptotics and typicality of sequential generalized measurements

Wen-Long Ma,1,2,* Shu-Shen Li,1,2 and Ren-Bao Liu3

1 State Key Laboratory of Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China
2 College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 100049, China
3 Department of Physics, Centre for Quantum Coherence, and The Hong Kong Institute of Quantum Information Science and Technology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

(Dated: August 18, 2022)

The relation between projective measurements and generalized quantum measurements is a fundamental problem in quantum physics, and clarifying this issue is also important to quantum technologies. While it has been intuitively known that projective measurements can be constructed from sequential generalized or weak measurements, there is still lack of a proof of this hypothesis in general cases. Here we rigorously prove it from the perspective of quantum channels. We show that projective measurements naturally arise from sequential generalized measurements in the asymptotic limit. Specifically, a selective projective measurement arises from a set of typical sequences of sequential generalized measurements. We provide an explicit scheme to construct projective measurements of a quantum system with sequential generalized quantum measurements. Remarkably, a single ancilla qubit is sufficient to mediate a sequential weak measurement for constructing arbitrary projective measurements of a generic system.

Quantum measurements retrieve classical information from quantum states [1, 2], and are particularly important to quantum technologies [3]. The traditional description of measurement in quantum mechanics is through projective measurements (PMs) of observables represented by Hermitian operators [4]. Measuring an observable corresponds to statistically projecting the quantum state to one of the orthogonal eigenspaces of this observable. PMs appear most commonly in quantum foundation and quantum information theory, and are widely useful for initialization and readout of quantum systems in quantum technologies [5–11].

Recent quantum theories show that there exist more general quantum measurements, called generalized measurements or positive-operator-valued measures (POVMs) [12–16]. POVMs can outperform PMs in many tasks in quantum technologies, such as quantum tomography [17] and quantum state discrimination or estimation [18, 19]. Moreover, continuous or sequential POVMs can be exploited for monitoring and maneuvering quantum evolutions [20–29]. In particular, weak measurements can extract partial information without projections, and therefore can help realize optimal qubit tomography [30], reconcile measurement incompatibility [31] and extract arbitrary bath correlations [32–34].

Substantial efforts have been devoted to illustrating the relation between PMs and POVMs. A celebrated result is Naimark’s theorem [4], stating that any POVM can be implemented as a PM on an enlarged Hilbert space. POVMs can also be simulated by PMs with classical randomness or postselection [35, 36]. In the opposite direction, it has been argued that sequential weak measurements can generate PMs by analysing the gradual state collapse [37–39] or the statistics of measurement results [40–42]. However, to our knowledge, the general relation between PMs and sequential POVMs still remains elusive.

In this paper, we rigorously prove that PMs emerge from sequential POVMs in the asymptotic limit, when the measurement operators are normal operators and commuting with each other. The proof is based on the observation that projective subspaces are fixed points of the quantum channels for such POVMs. Moreover, from the theory of classical typicality, we find that different selective PMs arise from different sets of typical sequences of sequential selective POVMs. These results completely characterize the structures of sequential commuting POVMs with normal measurement operators. We further present a general scheme to realize such POVMs with a single qubit ancilla, and show that sequential POVMs can be used to projectively measure an arbitrary finite-dimensional quantum system. An example of the scheme is also provided to measure the modular excitation numbers of an infinite-dimensional bosonic mode with an ancilla qubit.

**POVMs and quantum channels.** Consider a $d$-level quantum system under non-selective POVMs with the measurement operators $\{M_{\alpha}\}_{\alpha=1}^r$. Its evolution is described by the completely positive and trace-preserving (CPTP) quantum maps in the Kraus representation [12],

$$\Phi(\rho) = \sum_{\alpha=1}^r M_{\alpha}\rho M_{\alpha}^\dagger,$$  

(1)

where $M_{\alpha} = M_{\alpha}(\cdot)M_{\alpha}^\dagger$ is a superoperator in the Hilbert space representing a selective POVM with the $\alpha$th outcome, and $\sum_{\alpha=1}^r M_{\alpha}^\dagger M_{\alpha} = \mathbb{I}$ being the identity operator.
ensures the trace-preserving property of such maps. The CPTP maps are also called quantum operations or quantum channels [43, 44].

Quantum channels have convenient matrix representations in the Hilbert-Schmidt (HS) space of the quantum system [45, 46]. While the density matrices are operators in the Hilbert space with an orthonormal basis \( \{ |i\rangle \}_{i=1}^{d} \), they are turned into vectors in the HS space, i.e., \( \rho = \sum_{i,j=1}^{d} \rho_{ij} |i\rangle \langle j| \leftrightarrow |\rho\rangle \). Such that \( X \rho Y \leftrightarrow X \otimes Y^T |\rho\rangle \) with \( X, Y \) being operators in the Hilbert space and \( Y^T \) the transpose of \( Y \). The inner product in the HS space is defined as \( \langle \sigma | \rho \rangle = Tr[\sigma^T \rho] \). Then non-selective POVMs act as the following linear operators in the HS space,

\[
\hat{\Phi} |\rho\rangle = \sum_{\alpha=1}^{r} \tilde{M}_{\alpha} |\rho\rangle = \sum_{\alpha=1}^{r} M_{\alpha} \otimes M_{\alpha}^* |\rho\rangle,
\]

where \( M_{\alpha}^* \) is the complex conjugate of \( M_{\alpha} \). Note that we add hats for operators in the HS space, to distinguish them from the corresponding superoperators in the Hilbert space. With the HS space, the probability to get the \( \alpha \)th outcome is \( \langle \tilde{M}_{\alpha} |\rho\rangle = Tr(M_{\alpha} \rho M_{\alpha}^*) \).

**POVMs with normal and commuting measurement operators.** We assume that the set of measurement operators \( \{ M_{\alpha} \}_{\alpha=1}^{r} \) are all normal operators and commuting with each other, i.e., \( [M_{\alpha}, M_{\beta}] = [M_{\alpha}, M_{\delta}] = 0 \) for all integers \( \alpha, \beta, \delta \in [1, r] \), such that \( \{ M_{\alpha} \}_{\alpha=1}^{r} \) can be simultaneously diagonalized in an orthonormal eigenbasis of the quantum system [47],

\[
\begin{bmatrix}
M_1 \\
\vdots \\
M_r
\end{bmatrix} =
\begin{bmatrix}
c_{11} & \cdots & c_{1d} \\
\vdots & \ddots & \vdots \\
c_{r1} & \cdots & c_{rd}
\end{bmatrix}
\begin{bmatrix}
|1\rangle \\
\vdots \\
|d\rangle
\end{bmatrix}.
\]

(3)

This can be simply denoted as \( \mathbf{M} = \mathbf{CP} \), where \( \mathbf{M} = [M_1, \ldots, M_r]^T \), \( \mathbf{P} = [|1\rangle, \ldots, |d\rangle]^T \), and \( \mathbf{C} \) is a \( r \times d \) complex matrix (\( r \) and \( d \) are generally different).

We partition \( \mathbf{C} \) according to its columns as \( [\mathbf{c}_1, \cdots, \mathbf{c}_d] \), then \( \| \mathbf{c}_j \|^2 = \mathbf{c}_j^T \mathbf{c}_j = 1 \) for any integer \( j \in [1, d] \) due to \( \mathbf{M}^T \mathbf{M} = \sum_{\alpha=1}^{r} M_{\alpha}^T M_{\alpha} = \mathbf{I} \), and \( \{ \mathbf{c}_j \}_{j=1}^{d} \) is a set of unit vectors in a \( r \)-dimensional complex vector space, with \( j \) corresponding to the basis state \( |j\rangle \). Note that these unit vectors are not necessarily orthogonal to each other [46]. For a specific POVM, the measurement operators are not unique, since we can define a new set of measurement operators by \( \mathbf{M}' = \mathbf{T} \mathbf{M} \) with \( \mathbf{T} \) being a \( r \times r \) unitary matrix, which satisfy \( \mathbf{M}'^T \mathbf{M}' = \mathbf{I} \) and also characterize the same quantum channel.

The non-selective POVMs above act as diagonal operators in the HS space,

\[
\hat{\Phi} = \sum_{i,j=1}^{d} \mathbf{c}_j^T \mathbf{c}_i |ij\rangle \langle ij|,
\]

(4)

where \( \{ |ij\rangle \}_{i,j=1}^{d} \) are the eigenvectors (eigenmatrices in the Hilbert space) of \( \hat{\Phi} \) with the corresponding eigenvalues \( \{ \mathbf{c}_j^T \mathbf{c}_i \}_{i,j=1}^{d} \). Since \( |\mathbf{c}_j^T \mathbf{c}_i | \leq 1 \) (due to the Cauchy-Schwarz inequality) with equality if and only if \( \mathbf{c}_j = e^{i\varphi} \mathbf{c}_i \) for some real \( \varphi \), all the eigenvalues of \( \hat{\Phi} \) lie within the unit disk of the complex plane. The eigenvectors with eigenvalue 1 are called fixed points [44], and those with eigenvalues \( e^{i\varphi} \) with \( \varphi \neq 0 \) are rotating points. Obviously the fixed points are \( \{ |ij\rangle \}_{i,j=1}^{d} \), and the rotating points are \( \{ |ij\rangle \}_{i,j=1}^{d} \), but no rotating points. As another example, consider \( \{ \mathbf{c}_j \}_{j=1}^{d} \) as a set of orthonormal vectors, then the channel is \( \hat{\Phi} = \sum_{j=1}^{d} |jj\rangle \langle jj| \), representing PMs with rank-1 projectors (von Neumann measurements), \( \Phi(\cdot) = \sum_{j=1}^{d} |jj\rangle \langle j(\cdot) |j\rangle \). This channel has only fixed points but no rotating points. As another example, consider \( \{ \mathbf{c}_j \}_{j=1}^{d} = \{ e^{i\varphi_j} \}_{j=1}^{d} \), then \( \hat{\Phi} = \sum_{j=1}^{d} e^{i(\varphi_i - \varphi_j)} |ij\rangle \langle ij| \) is a unitary channel \( \Phi(\cdot) = U(\cdot)U^\dagger \) with \( U = \sum_{j=1}^{d} e^{i\varphi_j} |j\rangle \langle j| \). For the unitary channel, \( |ij\rangle \) is a fixed point if \( i = j \) or \( \varphi_i = \varphi_j \), and a rotating point if \( \varphi_i \neq \varphi_j \).
For general cases, we divide the index set $A = \{1, \cdots, d\}$ into $s (\leq d)$ disjoint subsets $A_1, \cdots, A_s$, with the corresponding cardinalities (number of elements) being $d_1, \cdots, d_s$, satisfying $\sum_{i=1}^{s} d_i = d$. Then divide the set of unit vectors $C = \{c_j\}_{j=1}^{d}$ into $s$ disjoint subsets $C_1, \cdots, C_s$, with $C_k = \{c_j | j \in A_k\}$. This division should ensure that the unit vectors in each subset are the same up to some phase factors but are different from any other unit vectors in other subsets, i.e., $C_k = \{\tilde{c}_ke^{i\varphi_j} | j \in A_k\}$ but $\tilde{c}_p \neq \tilde{c}_q e^{i\varphi_p}$ for any $\varphi$ and $p, q \in [1, s]$. This implies that $(|ij\rangle)$ with $i, j \in A_k$ is either a fixed point ($\varphi_i = \varphi_j$) or a rotating point ($\varphi_i \neq \varphi_j$).

The division of the index set also partitions the Hilbert space $H$ of the quantum system into the direct sum of $s$ subspaces, $H = H_1 \oplus \cdots \oplus H_s$, where $H_k = \{\text{Span}(|ij\rangle) | j \in A_k\}$ with rank-$d_k$ projection $P_k = \sum_{j \in A_k} |j\rangle\langle j|$. Thus the measurement operators in Eq. (3) can be written in a compact matrix form, $M = CP$, where $C = [c_1, \cdots, c_s]$ and $P = [P_1, \cdots, P_s]^T$ with $P_k = \sum_{j \in A_k} e^{i\varphi_j} |j\rangle\langle j|$. Note that $P_k$ is either a projection operator or a unitary operator in $H_k$, satisfying $P_k^2 = \delta_{kk} P_k$ and $\sum_{k=1}^s P_k = I$. Such a compact form of $M$ allows us to extend the above formulation to infinite-dimensional quantum systems [46], if we divide the identity operator into a finite set of orthogonal projections.

Asymptotics of sequential POVMs. Sequential POVMs correspond to sequential applications of the quantum channel $\hat{\Phi}$ [Fig. 1(a)]. Previous works have studied the asymptotic behaviors of sequential general quantum channels [48–51], mostly trying to find which information from an initial state can be preserved during the process.

For the channel in Eq. (4), as the number of applications $m$ increases, the projections to eigenvectors with eigenvalues lying in the interior of the unit disk $|\langle \tilde{c}_j | c_i \rangle| < 1$ gradually vanish, while the projections to eigenvectors with eigenvalues on the unit circle $|\langle \tilde{c}_j | c_i \rangle| = 1$ remain unchanged or change by some phase factors. So sequential POVMs tend to preserve the quantum coherence within subspaces $\{H_k\}_{k=1}^s$ but diminish the coherence between different subspaces. First assume that the elements in each $C_k$ are all the same, i.e., $\varphi_j = 0$ for all $j \in [1, d]$. In the asymptotic limit of large $m$,

$$\lim_{m \to \infty} \hat{\Phi}_m^m = \sum_{k=1}^s \sum_{i,j \in A_k} |ij\rangle\langle ij| = \sum_{k=1}^s \hat{P}_k,$$

(5)

which represents non-selective PMs. Then consider the possible phase factors in $C_k = \{\tilde{c}_ke^{i\varphi_j} | j \in A_k\}$, each application of $\hat{\Phi}$ produces a unitary operation in the Hilbert subspace $H_k$, i.e., $P_k(\cdot)P_k$ in the former case should be replaced by $\tilde{P}_k^m(\cdot)(\tilde{P}_k^m)^\dagger$ for example, if $C_k = \{c_i, c_j\} = \{\tilde{c}_ke^{i\varphi}, c_k e^{i\varphi}\}$, then $\tilde{P}_k = e^{i\varphi} |i\rangle\langle i| + e^{i\varphi} |j\rangle\langle j|$.

Typicality of sequential POVMs. Now that sequential non-selective POVMs produce projections (or unitary operations in the projected subspaces) in the asymptotic limit, we further ask which sequences of sequential selective POVMs produce a specific projection. This problem can be perfectly solved by the theory of classical typicality [52–56]. Classical typicality mainly cares about the following problem: if a random variable takes $r$ different values with the probability distribution $(p_1, \cdots, p_r)$, generate $m$ independent realizations of this variable and find the statistical distributions of the event sequences with $(m_1/m, \cdots, m_r/m)$, where $m_i$ is the number of the occurrences of the $i$th value. For infinitely large $m$, the event sequences that are overwhelmingly likely to occur are the set of typical sequences with $(p_1, \cdots, p_r)$. By defining sequences of selective POVMs as a product of superoperators [Fig. 1(a)], we find that the asymptotic projections are induced by the sets of typical sequences of selective POVMs.

Since $\hat{\Phi} = \sum_{\alpha=1}^r \hat{M}_\alpha$ and $[\hat{M}_\alpha, \hat{M}_\beta] = 0$ for $\alpha, \beta \in [1, r]$, we can expand $\Phi^m$ according to the multinomial theorem,

$$\Phi^m = \sum_{\{F\}} \frac{m!}{(m_{f_1})! \cdots (m_{f_r})!} \hat{M}_{f_1}^{m_{f_1}} \cdots \hat{M}_{f_r}^{m_{f_r}},$$

(6)

where $F = (f_1, \cdots, f_r)$ with $f_i \in [0, 1]$ (also a rational number with denominator $m$) satisfying $\sum_{i=1}^r f_i = 1$, and the summation is over all distributions $\{F\}$ in a ($r - 1$)-dimensional probability space. For large $m$, $\Phi^m$ can be approximated by its projections in the asymptotic subspaces [46],

$$\hat{\Phi}_m^m \approx \sum_{k=1}^s \hat{P}_k \hat{\Phi}_m^m \hat{P}_k \approx \sum_{k=1}^s e^{-mS(F||F_k)} \hat{P}_k^m,$$

(7)

where $F_k = (f_{k1}, \cdots, f_{kr}) = (|c_{k1}|^2, \cdots, |c_{kr}|^2)$ with $c_{k1}, \cdots, c_{kr}$ being entries of $c_k$ satisfying $\sum_{i=1}^r |c_{ki}|^2 = 1$, and $S(F||F_k) = \sum_{i=1}^r f_i \ln(f_i/f_{ki})$ is the relative entropy between $F$ and $F_k$ (the derivation above uses Stirling’s formula $\ln m! \approx m \ln m - m$ for large $m$). $S(F||F_k)$ takes the minimum when $F = F_k$, so for infinite large $m$, $\{F_k\}_{k=1}^s$ represent $s$ sets of ideal typical sequences of POVMs leading to the projections $\{\hat{P}_k^m\}_{k=1}^s$ correspondingly [Fig. 1(b)].

For large but finite $m$, the distributions of POVM sequences for $\hat{P}_k$ are concentrated around $F_k$, so $S(F||F_k) \approx \sum_{i=1}^r (f_i - f_{ki})^2/(2f_{ki})$. Then Eq. (7) represents the summation of $s$ Gaussians around $F_1, \cdots, F_s$, with integration of the $k$th Gaussian over the whole probability space giving rise to $\hat{P}_k$. For any two Gaussians around $F_j$ and $F_k$, they are well separated if the distance between $F_j$ and $F_k$ is larger than the sum of the respective Gaussian half widths. This requires $m > 2 |\ln(n)| + |\sum_{i=1}^r (f_{ji} - f_{ki})^2/(2f_{ki})|^{-1/2} \approx |\sum_{i=1}^r (f_{ji} - f_{ki})^2|^{-1/2} \approx 1$.


which can be written in the Stinespring representation as [50]

\[
\Phi(\rho) = \text{Tr}_\alpha[U(t)(\rho_\alpha \otimes \rho)U^\dagger(t)],
\]

where \( U(t) = T e^{-i\sigma \oint_0^t B(t')dt'} \) with \( T \) being the time-ordering operator, \( \rho_\alpha = |\psi\rangle\langle\psi| \) is the initial state of the ancilla, \( \rho \) denotes the density matrix of the target system, and \( \text{Tr}_\alpha \) denotes the partial trace over the ancilla. With an orthonormal ancilla basis \( \{|\psi_\alpha\rangle\} \), we obtain the Kraus representation of the quantum channels, \( \Phi(\rho) = \sum_{\alpha \in \{+,-\}} M_{\alpha\beta} \rho M_{\alpha\beta}^\dagger \) with \( M_{\alpha\beta} = \langle \psi_\alpha|U(t)|\psi_\beta \rangle \) (note that we add subscripts to the kets only when representing matrix elements or inner products with respect to the ancilla states). With another orthonormal basis \( \{|T|_+\rangle, |T|_-\rangle\} \) with \( T \) being a unitary operator for the ancilla, the measurement operators become \( \{M'_{\alpha\beta}\} \) with \( M'_{\alpha\beta} = \sum_{\beta \in \{+,-\}} T_{\alpha\beta} M_\beta \), while the quantum channels remain unchanged.

We expand \( U(t) \) in the ancilla eigenbasis \( \{|+\rangle_\alpha, |-\rangle_\alpha\} \) of \( \sigma_z \) as

\[
U(t) = |+\rangle_\alpha \langle +| \otimes U_\perp(t) + |-\rangle_\alpha \langle -| \otimes U_- (t),
\]

where \( U_\perp(t) = T e^{i\sigma \oint_0^t B(t')dt'} \). If \( U_\perp(t) \) is exactly equal to or well approximated by its first-order Magnus expansion [60], i.e., \( U_\perp(t) = e^{i\sigma \oint_0^t B(t')dt'} \), then \( U_\perp(t) = U_\perp^\dagger(t) \) and \( |U_\perp(t), U_- (t)\rangle = 0 \), so \( U_\perp(t) \) and \( U_- (t) \) can be simultaneously diagonalized as \( U_\pm(t) = \sum_j e^{i\phi_j} |j\rangle \langle j| \). So the measurement operators are

\[
M_\pm = \sum_{j=1}^d (|\psi_{+}\rangle \langle \psi_{+}| \cos \omega_j + i|\psi_{-}\rangle \langle \psi_{-}| \sin \omega_j) |j\rangle \langle j|.
\]

As a special case, take \( |\psi_\alpha\rangle = R_{\phi_\alpha}(\frac{\pi}{2}) |+_\alpha \rangle \) and \( |\psi_\pm\rangle = R_{\phi}_\pm(\frac{\pi}{2}) |\pm\rangle \) with \( R_{\phi}(\theta) = e^{-i\phi(\cos \phi_\sigma + \sin \phi_\sigma \theta)/2} \), then

\[
\begin{bmatrix}
M_+ \\
M_-
\end{bmatrix} = \begin{bmatrix}
e^{i\omega_1} & e^{i(\Delta \phi - \omega_1)} & \ldots & e^{i\omega_d} & e^{i(\Delta \phi - \omega_d)} \\
e^{i\omega_1} & e^{i(\Delta \phi - \omega_1)} & \ldots & e^{i\omega_d} & e^{i(\Delta \phi - \omega_d)}
\end{bmatrix} \frac{\mathbf{P}}{2}
\]

(10)

where \( \Delta \phi = \phi_1 - \phi_2 \). Each round of such POVMs corresponds to a three-step physical process [Fig. 2(a)]: (1) the ancilla starts from \( |+_\alpha \rangle \) and is rotated by \( R_{\phi_\alpha}(\frac{\pi}{2}) \); (2) let the ancilla and target systems evolve under \( H(t) \) for time \( t \); (3) finally rotate the ancilla by \( R_{\phi_\pm}(\frac{\pi}{2}) \) and make a PM of the ancilla in the basis \( \{|+_\alpha\rangle, -\rangle_\alpha\} \). Similar schemes have been designed to realize single-shot readouts of nuclear spins-1/2 in diamond [41], but here we show this scheme can be extended to perform PMs of a generic system.

Since the POVMs have only two outcomes, the measurement results are solely determined by the polarization \( \Delta f_j = \langle m_- - m_+ \rangle / m \) [40], with \( m_+ / m_- \) being the number of outcome +/− in \( m \) sequential measurements. For the spectra \( \{e^{\pm i\omega_j}\} \) of \( U_\perp(t) \), calculate \( \Delta f_j = \cos(2\omega_j - \Delta \phi) \) for all \( j \in [1,d] \). Weak measurement corresponds to the regime \( |\Delta f_j| \ll 1 \). If \( \Delta f_j \neq \Delta f_k \) for any \( j, k \in [1,d] \) and \( j \neq k \), sequential POVMs produce von-Neumann measurements of the target system, with the rank-1 projection \( P_j = |j\rangle \langle j| \) corresponding to

FIG. 2. (a) Quantum circuit diagram to realize sequential POVMs on the target system with PMs of an ancilla qubit. (b) Diagrams of eigenvalues of \( U_\perp(t) = e^{i\pi \alpha \oint \sigma A dt/2} \) in the complex unit circle to detect the \( k \) mod \( N \) excitation numbers of a bosonic mode for \( t = 2\pi/(N\chi) \) and \( N = 2, 3 \), \( f_k(t)^2 / f_k(1)^{-1/2} \) [46, 57], where \( \eta \) is the ratio of the minimum hight to the maximum hight within the Gaussian width. If all the Gaussians are well separated, integration of the POVM sequences within a small neighborhood around \( F_j \) can approximate \( \mathcal{P}_j \) up to arbitrary small error as \( m \) increases (see the Supplementary Material [46] for the error rates with finite \( m \)).

It may happen that two Gaussians coincide around \( F_j = F_k \) but \( \mathbf{c}_j \neq \mathbf{c}_k \), i.e., partial elements of \( \mathbf{c}_j \) and \( \mathbf{c}_k \) differ by some phase factors. Since \( |\mathbf{c}_j^\dagger \mathbf{c}_k| < 1 \), the coinciding Gaussians actually correspond to different projections, and the POVM sequences around \( F_j \) approximately produce \( \mathcal{P}_j \) + \( \mathcal{P}_k \). To realize selective projections, we can get a new set of measurement operators by a unitary transformation, thus creating different typical sequences of POVMs for \( \mathcal{P}_j \) and \( \mathcal{P}_k \).

Physical realization. We present a general physical model to perform PMs on a \( d \)-level quantum system (called target system) with sequential POVMs. Without loss of generality, we assume that the POVMs are realized by PMs of an ancilla qubit. The coupling Hamiltonian of the composite system (including the ancilla and target systems) is in the pure-dephasing form [58]

\[
H(t) = \sigma_z \otimes B(t),
\]

where \( \sigma_z \) is the Pauli-\( i \) operator of the ancilla qubit (\( i = x, y, z \)) and \( B(t) \) is a time-dependent Hermitian operator of the target system (the time-dependence of \( B(t) \) is due to being in some interaction picture or external drivings).

The dynamics of the composite system induces a general class of quantum channels on the target system,
typical POVM sequences with $\Delta f_j$. If $\Delta f_j = \Delta f_k$, then either (I) $\omega_j + \omega_k = N\phi + n\pi$ or (II) $\omega_j - \omega_k = n\pi$ with $n$ being integers. In case-I, the typical POVM sequences for $P_j$ and $P_k$ are the same, but selective projections can still be achieved by choosing a different $\Delta \phi'$. In case-II, the typical POVM sequences with $\Delta f_j$ induce the operation $P_j + (-1)^n P_k$.

Example: Modular excitation number measurements of bosonic modes. As an example, we present a protocol to modulate the modular excitation numbers of a bosonic mode with an ancilla qubit. The ancilla is dispersive coupled to a bosonic mode with $H = -\chi a^\dagger a^2/2$, where $a$ ($a^\dagger$) is the annihilation (creation) operator of the bosonic mode and $\chi$ is the dispersive coupling strength. The dispersive coupling arises naturally from the Jaynes-Cummings coupling in cavity quantum electrodynamics (QED) [61] and circuit QED [62] when the detuning between the ancilla and the bosonic mode is much larger than the coupling strength.

We construct the projectors into the sets of bosonic Fock states with modular excitation number $l$ mod $2N$, $P_{2N}^k = \sum_{\ell = 0}^{2N-1} |2\ell\rangle\langle 2\ell |$, with $l \in \{0, 1, \ldots, 2N - 1\}$ and $N$ being any positive integer. With the scheme below Eq. (10) and the evolution time $t = 2\pi/(N\chi)$, $U_{\pm}(t) = e^{\pm i x a^\dagger /2} = \sum_{k=0}^{N-1} e^{\pm i k\pi/N}(P_k^k-P_{2N}^{k+N})$, i.e. the eigenvalues of $U_{\pm}(t)$ divides the complex unit circle into $2N$ equal pieces [Fig. 2(b)]. The measurement operators are $M_{\pm} = \sum_{k=0}^{N-1}(e^{i k\pi/N} \mp e^{i (\Delta \phi - k\pi/2N)}(P_k^k-P_{2N}^{k+N})$, and the polarization $\Delta f_k = \cos(2k\pi/N - \Delta \phi)$. We can tune $\Delta \phi$ so that $\Delta f_k$ is maximally distinguishable for different $k \in [0, N - 1]$. For $N = 1$, $\Delta \phi = 0$ is optimal as $\Delta f_0 = -\Delta f_1 = 1$; while for $N \geq 2$, we can choose $\Delta \phi = \pi/(2N)$ so that $\Delta f_k = \cos[(2k-1)/2]\pi/N$. Then for a large and even $m$, sequential POVMs induce the $k$ mod $N$ excitation number measurement of the bosonic mode. The modular excitation numbers are the error syndromes of rotation-symmetric error correction codes of bosonic modes [63], such as cat codes [64–67] and binomial codes [68]. So this protocol is useful for quantum non-demolition measurements in bosonic quantum information processing [69–72], especially for tracking the error syndromes of high-order bosonic error correction codes [73–75].

Summary. We have revealed the elegant structures of sequential POVMs by studying their asymptotic behaviors and typical sequences. We rigorously prove that PMs emerge from sequential POVMs when the measurement operators are normal and commuting with each other. Each selective projective measurement comes from a set of typical sequences of sequential POVMs, which is determined solely by the structures of the measurement operators. While the POVMs here are restricted to have normal and commuting measurement operators, they describe a large class of quantum channels on a quantum system induced by a pure-dephasing coupling between this system and an ancilla system. For future works, it will be interesting to relax this restriction, and study the asymptotics and typicality of sequential POVMs with general measurement operators.

We acknowledge support from the Startup Foundation of Institute of Semiconductors, Chinese Academy of Sciences (No. E05EBB11), and National Natural Science Foundation of China (No. 12174379). R.B.L. was supported by the Hong Kong Research Grants Council – General Research Fund Project 14300119.

* wenlongma@semi.ac.cn

[1] H. M. Wiseman and G. J. Milburn, Quantum Measurement and Control (Cambridge University Press, Cambridge, England, 2010).
[2] K. Jacobs, Quantum Measurement Theory and Its Applications (Cambridge University Press, Cambridge, England, 2014).
[3] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, England, 2010).
[4] A. Peres, Quantum Theory: Concepts and Methods (Springer Science & Business Media, New York, 2006), Vol. 57.
[5] J. M. Eizerman, R. Hanson, L. H. W. Van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven, Single-Shot Read-out of an Individual Electron Spin in a Quantum Dot, Nature 430, 431 (2004).
[6] A. N. Vamivakas, C.-Y. Lu, C. Matthiesen, Y. Zhao, S. Fält, A. Badolato, and M. Atatüre, Observation of spin-dependent quantum jumps via quantum dot resonance fluorescence, Nature 467, 297 (2010).
[7] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko, Single-shot readout of a single nuclear spin, Science 329, 542 (2010).
[8] A. Morello, J. J. Pla, F. A. Zwanenburg, K. W. Chan, K. Y. Tan, H. Huebl, M. Möttönen, C. D. Nugroho, C. Yang, J. A. van Donkelaar, A. D. C. Alves, D. N. Jamieson, C. C. Escott, L. C. L. Hollenberg, R. G. Clark, and A. S. Dzurak, Single-shot readout of an electron spin in silicon, Nature 467, 687 (2010).
[9] L. Jiang, J. S. Hodges, J. R. Maze, P. Maurer, J. M. Taylor, D. G. Cory, P. R. Hemmer, R. L. Walsworth, A. Yacoby, A. S. Zibrov, and M. D. Lukin, Repetitive readout of a single electronic spin via quantum logic with nuclear spin ancillae, Science 326, 267 (2009).
[10] T. Nakajima et al., Robust Single-Shot Spin Measurement with 99.5% Fidelity in a Quantum Dot Array, Phys. Rev. Lett. 119, 017701 (2017).
[11] A. West et al., Gate-Based Single-Shot Readout of Spins in Silicon, Nat. Nanotechnol. 14, 437 (2019).
[12] K. Kraus, States, Effects, and Operations: Fundamental Notions of Quantum Theory, Lecture Notes in Physics Vol. 190 (Springer-Verlag, Berlin, 1983).
[13] E. Andersson and D. K. L. Oi, Binary search trees for generalized measurements, Phys. Rev. A 77, 052108 (2008).
[14] Y. -H. Chen and T. A. Brun, Decomposing qubit positive-operator-valued measurements into continuous
destructive weak measurements, Phys. Rev. A 98, 062113 (2018).
[15] Y. -H. Chen and T. A. Brun, Qubit positive-operator-valued measurements by destructive weak measurements, Phys. Rev. A 99, 062121 (2019).
[16] Y. W. Cheong and S. -W. Lee, Balance between Information Gain and Reversibility in Weak Measurement, Phys. Rev. Lett. 109, 150402 (2012).
[17] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. 45, 2171 (2004).
[18] J. A. Bergou, Discrimination of quantum states, J. Mod. Opt. 57, 160 (2010).
[19] R. Derka, V. Bužek, and E. A. Kert, Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement, Phys. Rev. Lett. 80, 1571 (2001).
[20] K. Jacobs and D. A. Steck, A straightforward introduction to continuous quantum measurement, Contemp. Phys. 47, 279 (2006).
[21] S. A. Gurvitz, Measurements with a noninvasive detector and dephasing mechanism, Phys. Rev. B 56, 15215 (1997).
[22] S. Ashhab, J. Q. You, and F. Nori, The information about the state of a qubit gained by a weakly coupled detector, New J. Phys. 11, 083017 (2009).
[23] A. N. Korotkov, Sequential quantum evolution of a qubit state due to continuous measurement, Phys. Rev. B 63, 115403 (2001).
[24] M. S. Blek, C. Bonato, M. L. Markham, D. J. Twitchen, V. V. Dobrovitski, and R. Hanson, Manipulating a qubit through the backaction of sequential partial measurements and real-time feedback, Nat. Phys. 10, 189 (2014).
[25] A. N. Jordan and A. N. Korotkov, Qubit feedback and control with kicked quantum nondemolition measurements: A quantum Bayesian analysis, Phys. Rev. B 74, 085307 (2006).
[26] A. Chantasri, J. Dressel, and A. N. Jordan, Action principle for continuous quantum measurement, Phys. Rev. A 88, 042110 (2013).
[27] A. Chantasri and A. N. Jordan, Stochastic path-integral formalism for continuous quantum measurement, Phys. Rev. A 92, 032125 (2015).
[28] C. Presilla, R. Onofrio, and U. Tambini, Measurement quantum mechanics and experiments on quantum Zeno effect, Ann. Phys. 248, 95 (1996).
[29] L. Díósi, Structural features of sequential weak measurements, Phys. Rev. A 94, 010103 (2016).
[30] E. Shojaei, C. S. Jackson, C. A. Riofrío, A. Kalev, and I. H. Deutsch, Optimal Pure-State Qubit Tomography via Sequential Weak Measurements, Phys. Rev. A 92, 032125 (2015).
[31] J. T. Monroe, N. Yunger Halpern, T. Lee, and K. W. Murch, Weak Measurement of a Superconducting Qubit Reconciles Incompatible Operators, Phys. Rev. Lett. 126, 100403 (2021).
[32] P. Wang, C. Chen, X. Peng, J. Wrachtrup, and R. B. Liu, Characterization of Arbitrary-Order Correlations in Quantum Baths by Weak Measurement, Phys. Rev. Lett. 123, 50603 (2019).
[33] M. Pfleider et al., High-resolution spectroscopy of single nuclear spins via sequential weak measurements, Nat. Commun. 10, 594 (2019).
[34] Z. Wu, P. Wang, T. Wang, Y. Li, R. Liu, Y. Chen, X. Peng, R.-B. Liu, and J. Du, Detection of Arbitrary Quantum Correlations via Synthesized Quantum Channels, arXiv: 2206.05883 (2022).
[35] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, Phys. Rev. Lett. 119, 190501 (2017).
[36] M. Oszmaniec, F. B. Maciejewski, and Z. Puchala, Simulating all quantum measurements using only projective measurements and postselection, Phys. Rev. A 100, 01235 (2019).
[37] T. A. Brun, A simple model of quantum trajectories, Am. J. Phys. 70, 719 (2002).
[38] D. Lidar and T. D. Brun, Quantum Error Correction (Cambridge University Press, Cambridge, England, 2013).
[39] O. Oreshkov and T. A. Brun, Weak Measurements Are Universal, Phys. Rev. Lett. 95, 110409 (2005).
[40] W. -L. Ma, P. Wang, W. -H. Leong, and R. -B. Liu, Phase transitions in sequential weak measurements, Phys. Rev. A 98, 012117 (2018).
[41] G. -Q. Liu, J. Xing, W. -L. Ma, P. Wang, C. -H. Li, H. C. Po, Y. -R. Zhang, H. Fan, R. -B. Liu, and X. -Y. Pan, Single-Shot Readout of a Nuclear Spin Weakly Coupled to a Nitrogen-Vacancy Center at Room Temperature, Phys. Rev. Lett. 118, 150504 (2017).
[42] D. D. Bhaktavatsala Rao, S. Yang, S. Jesenski, E. Tekin, F. Kaiser, and J. Wrachtrup, Observation of nonclassical measurement statistics induced by a coherent spin environment, Phys. Rev. A 100, 22307 (2019).
[43] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Quantum channels and memory effects, Rev. Mod. Phys. 86, 1203 (2014).
[44] M. M. Wolf, Quantum Channels & Operations Guided Tour, 2010.
[45] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press, Cambridge, England, 2017).
[46] See Supplementary Material for details about the HS space and POVM, structural properties of POVMs with normal and commuting measurement operators, matrix representation of quantum channels for POVMs, derivations in typicality of sequential POVMs including the conditions and error rates for approximating sequential POVMs with PMs.
[47] S. R. García and R. A. Horn, A Second Course in Linear Algebra (Cambridge University Press, Cambridge, England, 2017).
[48] V. V. Albert, Asymptotics of quantum channels: conserved quantities, an adiabatic Limit, and matrix products, Quantum 3, 151 (2019).
[49] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, Ergodic and mixing quantum channels in finite dimensions, New J. Phys. 15, 073045 (2013).
[50] J. Novotná, J. Maryška, and J. Iex, Quantum Markov processes: From attractor structure to explicit forms of asymptotic states: Asymptotic dynamics of quantum Markov processes, Eur. Phys. J. Plus 133, 310 (2018).
[51] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, Information-Preserving Structures: A General Framework for Quantum Zero-Error Information, Phys. Rev. A 82, 062306 (2010).
[52] M. W. Wilde, Quantum Information Theory (Cambridge University Press, Cambridge, England, 2017).
[53] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley-Interscience, New York, New Jersey, 2006).
[54] P. Facchi, S. Pascazio, and F. V. Pepe, Quantum typicality and initial conditions, Phys. Ser. 90, 074057 (2015).
[55] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Canonical Typicality, Phys. Rev. Lett. 96, 050403 (2006).
[56] C. Bartsch and J. Gemmer, Dynamical Typicality of Quantum Expectation Values, Phys. Rev. Lett. 102, 110403 (2009).
[57] R. B. Liu, W. Yao, and L. J. Sham, Quantum Computing by Optical Control of Electron Spins, Adv. Phys. 59, 703 (2010).
[58] W. Yang, W. -L. Ma, and R. -B. Liu, Quantum Many-Body Theory for Electron Spin Decoherence in Nanoscale Nuclear Spin Baths, Rep. Prog. Phys. 80, 16001 (2017).
[59] W. F. Stinespring, Positive functions on C\(^*\)-algebras, Proc. Am. Math. Soc. 6, 211 (1955).
[60] W. -L. Ma and R. -B. Liu, Angstrom-Resolution Magnetic Resonance Imaging of Single Molecules via Wave-Function Fingerprints of Nuclear Spins, Phys. Rev. Appl. 6, 024019 (2016).
[61] J. M. Raimond, M. Brune, and S. Haroche, Colloquium: Manipulating Quantum Entanglement with Atoms and Photons in a Cavity, Rev. Mod. Phys. 73, 565 (2001).
[62] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit Quantum Electrodynamics, Rev. Mod. Phys. 83, 025005 (2021).
[63] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, Quantum Computing with Rotation-Symmetric Bosonic Codes, Phys. Rev. X 10, 11058 (2020).
[64] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, Hardware-Efficient Autonomous Quantum Memory Protection, Phys. Rev. Lett. 111, 120501 (2013).
[65] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, Dynamically Protected Cat-Qubits: A New Paradigm for Universal Quantum Computation, New J. Phys. 16, (2014).
[66] L. Li, C. L. Zou, V. V. Albert, S. Muralidharan, S. M. Girvin, and L. Jiang, Cat Codes with Optimal Decoherence Suppression for a Lossy Bosonic Channel, Phys. Rev. Lett. 119, 030502 (2017).
[67] M. Bergmann and P. Van Loock, Quantum Error Correction against Photon Loss Using Multicomponent Cat States, Phys. Rev. A 94, 042332 (2016).
[68] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New Class of Quantum Error-Correcting Codes for a Bosonic Mode, Phys. Rev. X 6, 031006 (2016).
[69] A. Blais, S. M. Girvin, and W. D. Oliver, Quantum Information Processing and Quantum Optics with Circuit Quantum Electrodynamics, Nat. Phys. 16, 247 (2020).
[70] W. Cai, Y. Ma, W. Wang, C.-L. Zou, and L. Sun, Bosonic Quantum Error Correction Codes in Superconducting Quantum Circuits, Fundam. Res. 1, 50 (2021).
[71] A. Joshi, K. Noh, and Y. Y. Gao, Quantum Information Processing with Bosonic Qubits in Circuit QED, Quantum Sci. Technol. 6, 033001 (2021).
[72] W. -L. Ma, S. Puri, R. J. Schoelkopf, M. H. Devoret, S. M. Girvin, and L. Jiang, Quantum Control of Bosonic Modes with Superconducting Circuits, Sci. Bull. 66, 1789 (2021).
[73] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. M. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, L. Frunzio, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Tracking Photon Jumps with Repeated Quantum Non-Demolition Parity Measurements, Nature 511, 444 (2014).
[74] N. Ołek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Extending the Lifetime of a Quantum Bit with Error Correction in Superconducting Circuits, Nature 536, 441 (2016).
[75] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C. L. Zou, S. M. Girvin, L. M. Duan, and L. Sun, Quantum Error Correction and Universal Gate Set Operation on a Binomial Bosonic Logical Qubit, Nat. Phys. 15, 503 (2019).
Supplementary Material for “Asymptotics and typicality of sequential generalized measurements”

Wen-Long Ma,1,2,* Shu-Shen Li,1,2 and Ren-Bao Liu3

1State Key Laboratory of Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China
2College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 100049, China
3Department of Physics, Centre for Quantum Coherence, and The Hong Kong Institute of Quantum Information Science and Technology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

(Dated: August 18, 2022)

* wenlongma@semi.ac.cn
In this Supplementary Material, we provide an introduction to the basic concepts and the details of derivations in the main text. In Sec. I, we briefly introduce the Hilbert-Schmidt (HS) space and positive-operator-valued measures (POVMs). Then we provide a systematic description of POVMs with normal and commuting measurement operators in Sec. II, and the matrix representation of quantum channels for such POVMs in Sec. III. In Sec. IV, we give the detailed derivations about determining the typical sequences of POVMs for a specific projective measurement (PM), the conditions to distinguish different typical POVM sequences, and the error rates in approximating PMs with sequential POVMs.

I. HILBERT-SCHMIDT SPACE AND POVM

For a $d$-dimensional quantum system, the space of operators form a linear vector space. This is easily seen if the $d \times d$ complex matrix of an operator $X$ in an orthonormal eigenbasis $\{|i\rangle\}_{i=1}^d$ is reshaped into a $d^2 \times 1$ column vector,

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dd} \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}, \quad \iff \quad |X\rangle = \begin{bmatrix} x_1^T \\ \vdots \\ x_d^T \end{bmatrix}, \quad (S1)$$

where $x_i$ is the $i$th row of $X$ with $i \in [1, d]$, and $x_i^T$ is the transpose of $x_i$. With Dirac notations, the matrix reshaping can be simply represented by $X = \sum_{i,j=1}^d x_{ij} |i\rangle \langle j| \equiv |X\rangle = \sum_{i,j=1}^d x_{ij} |i\rangle \langle j|)$. Then the ordinary scalar product between $|X\rangle$ and $|Y\rangle$ defines an inner product between $X$ and $Y$,

$$\langle Y | X \rangle = \sum_{i=1}^d x_i^* x_i^T = \sum_{i,j=1}^d |x_{ij}|^2 = \text{Tr}(Y^T X), \quad (S2)$$

which is the so-called Hilbert-Schmidt (HS) inner product. The HS space is the space of operators equipped with the HS inner product.

The density matrices of the quantum system, as the class of positive operators with trace one, are also vectors in the HS space. In the HS space, the trace one constraint of a density matrix $\rho$ is equivalent to $\langle \langle I | \rho \rangle \rangle = \text{Tr}(\rho) = 1$, with $I$ being the identity operator. Left and right multiplications of $\rho$ by operators $X$ and $Y$ corresponds to multiplying $|\rho\rangle$ with a $d^2 \times d^2$ matrix,

$$X \rho Y = \sum_{i,j=1}^d x_{ik} y_{lj} \rho_{kl} |i\rangle \langle j|, \quad \iff \quad X \otimes Y^T |\rho\rangle. \quad (S3)$$

So the operation $X \cdot Y$ as a superoperator in the Hilbert space is equivalent to a linear operator $X \otimes Y^T$ in the HS space.

For the quantum channel of a non-selective POVM with the measurement operators $\{M_\alpha\}_{\alpha=1}^r$, the transformation from its Kraus representation in the Hilbert space to the matrix representation in the HS space is given by

$$\Phi(\rho) = \sum_{\alpha=1}^r M_\alpha \rho M_\alpha^T, \quad \IFF \quad \hat{\Phi}(\rho) = \sum_{\alpha=1}^r \hat{M}_\alpha |\rho\rangle = \sum_{\alpha=1}^r M_\alpha \otimes M_\alpha^* |\rho\rangle. \quad (S4)$$

where $M_\alpha = M_\alpha (\cdot) M_\alpha^T$ is a superoperator in the Hilbert space representing a selective POVM with the $\alpha$th outcome, $\sum_{\alpha=1}^r M_\alpha^T M_\alpha = I$ ensures the trace-preserving property, $\hat{M}_\alpha = M_\alpha \otimes M_\alpha^*$ is an operator in the HS space corresponding to $M_\alpha$, and $M_\alpha^*$ is the complex conjugate of $M_\alpha$. For a selective POVM with the $\alpha$th outcome, the density matrix undergoes the following evolution,

$$\rho_\alpha = \frac{M_\alpha \rho M_\alpha^T}{p_\alpha}, \quad \IFF \quad |\rho_\alpha\rangle = \frac{\hat{M}_\alpha |\rho\rangle}{p_\alpha} = \frac{M_\alpha \otimes M_\alpha^* |\rho\rangle}{p_\alpha}, \quad (S5)$$

where $p_\alpha = \text{Tr}(M_\alpha \rho M_\alpha^T) = \langle \langle I | \hat{M}_\alpha |\rho\rangle \rangle$ is the probability to get the $\alpha$th outcome, satisfying $\sum_{\alpha=1}^r p_\alpha = 1$. 
II. POVMS WITH NORMAL AND COMMUTING MEASUREMENT OPERATORS

In this section, we provide a systematic description of POVMs with normal and commuting measurement operators \( \{M_\alpha\}_{\alpha=1}^r \). Since \( [M_\alpha, M_\beta] = [M_\alpha, M_\beta] = 0 \) for all integers \( \alpha, \beta \in [1, r] \), \( \{M_\alpha\}_{\alpha=1}^r \) can be simultaneously diagonalized in an orthonormal eigenbasis \( \{|i\rangle\}_{i=1}^d \) of the quantum system \([1]\),

\[
\begin{bmatrix}
M_1 \\
\vdots \\
M_r
\end{bmatrix} = \begin{bmatrix}
c_{11} & \cdots & c_{1d} \\
\vdots & \ddots & \vdots \\
c_{r1} & \cdots & c_{rd}
\end{bmatrix} \begin{bmatrix}
|1\rangle\langle 1| \\
\vdots \\
|d\rangle\langle d|
\end{bmatrix},
\tag{S6}
\]

which can be written in a matrix form as \( M = CP \), with the definitions below

\[
M = \begin{bmatrix}
M_1 \\
\vdots \\
M_r
\end{bmatrix}, \quad C = [c_1, \cdots, c_d] = \begin{bmatrix}
c_{11} & \cdots & c_{1d} \\
\vdots & \ddots & \vdots \\
c_{r1} & \cdots & c_{rd}
\end{bmatrix}, \quad P = \begin{bmatrix}
|1\rangle\langle 1| \\
\vdots \\
|d\rangle\langle d|
\end{bmatrix}
\tag{S7}
\]

where \( M \) is a \( r \times 1 \) column vector of operators, \( C \) is a \( r \times d \) complex matrix with \( c_i \) being its \( i \)th column, and \( P \) is a \( d \times 1 \) column vector of operators. Further define \( M_i = [M_i^1, \cdots, M_i^r] \) and \( P_i^j = |i\rangle\langle j| \), then

\[
M_i M = \sum_{i=1}^r M_i^j M_i = P_i^j P = \sum_{i=1}^d |i\rangle\langle i| = I. \tag{S8}
\]

which clearly shows \( c_i^j c_i = \sum_{j=1}^d |c_{ij}|^2 = 1 \) for any \( i \in [1, d] \), i.e., all the columns \( \{c_{ij}\}_{j=1}^d \) of \( C \) are unit vectors in a \( r \)-dimensional complex vector space. But these unit vectors are not necessarily orthogonal to each other. The reason is that entries of \( P \) are not real or complex numbers but rank-1 projectors \( \{|i\rangle\langle i|\}_{i=1}^d \), satisfying \( |i\rangle\langle i| |j\rangle = \delta_{ij} |i\rangle\langle j| \).

Now we take a closer look at the structures of matrix \( C \). Define the set of its column vectors as \( C = \{c_j\}_{j=1}^d \) with a index set \( A = \{1, \cdots, d\} \). Then divide \( C \) into \( s \) disjoint subsets \( C_1, \cdots, C_s \) with the corresponding index subsets \( A_1, \cdots, A_s \), where \( C_k = \{c_j \in A_k \} \) for any integer \( k \in [1, s] \). The cardinality of \( C_k \) and \( A_k \) is \( d_k \), satisfying \( \sum_{k=1}^s d_k = d \) and \( d_k \geq 1 \). This division should ensure that the unit vectors in each subset are the same up to some phase factors but are different from any other unit vectors in other subsets, i.e., \( C_k = \{\overline{c}_k e^{i\varphi} | j \in A_k \} \) but \( \overline{c}_k \neq \overline{c}_q e^{i\varphi} \) for any real \( \varphi \) and \( p, q \in [1, s] \). This means that we can always simultaneously reorder the columns of \( C \) and the entries of \( P \), and then realob the eigenbasis \( \{|i\rangle\}_{i=1}^d \), so that \( C \) is in the following canonical form,

\[
C = [\overline{c}_1 e^{i\varphi_1}, \cdots, \overline{c}_1 e^{i\varphi_{d_1}}, \overline{c}_2 e^{i\varphi_{d_1+1}}, \cdots, \overline{c}_2 e^{i\varphi_{d_1+d_2}}, \cdots, \overline{c}_s e^{i\varphi_{d-s+1}}, \cdots, \overline{c}_s e^{i\varphi_{d-1}}],
\tag{S9}
\]

and \( P \) remains unchanged. Then Eq. (S6) can be written in a more compact matrix form, \( M = \overline{C}P \), where \( \overline{C} = [\overline{c}_1, \cdots, \overline{c}_s] \) and \( \overline{P} = [\overline{P}_1, \cdots, \overline{P}_s]^T \) with \( \overline{P}_k = \sum_{j \in A_k} e^{i\varphi_j} |j\rangle \langle j| \). More explicitly,

\[
\begin{bmatrix}
M_1 \\
\vdots \\
M_r
\end{bmatrix} = \begin{bmatrix}
\overline{c}_{11} & \cdots & \overline{c}_{1s} \\
\vdots & \ddots & \vdots \\
\overline{c}_{r1} & \cdots & \overline{c}_{rs}
\end{bmatrix} \begin{bmatrix}
\overline{P}_1 \\
\vdots \\
\overline{P}_s
\end{bmatrix}.
\tag{S10}
\]

While Eq. (S6) mainly concerns about finite-dimensional quantum systems, Eq. (S10) can describe both finite- and infinite-dimensional systems. The key point is to first partition the identity operator \( I \) of a generic system into a set of orthogonal projections \( \{P_k\}_{k=1}^s \), which corresponds to partitioning the Hilbert space \( \mathcal{H} \) of the system into the direct sum of \( s \) subspaces, \( \mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_s \). Then \( \overline{P}_k \) is a unitary operator in subspace \( \mathcal{H}_k \) (with projection \( P_k \) as a special case), satisfying \( \overline{P}_k^\dagger \overline{P}_k = \delta_{kk}, P_k \) and \( \overline{P}_k^\dagger = \sum_{k=1}^s \overline{P}_k^\dagger \overline{P}_k = I \). Obviously, projective measurements and unitary evolutions are both special cases of Eq. (S10).

Moreover, for both Eq. (S6) and Eq. (S10), we can define a new set of measurement operators by \( M' = TM \) with \( T = [T_{\alpha\beta}] \) being a \( r \times r \) unitary matrix, which satisfies \( M'^\dagger M' = MT^\dagger TM = I \). \( M' \) and \( M \) also characterize the
We define a distribution $F$ same CPTP map, since

$$\sum_{\alpha=1}^{r} M'_{\alpha}(\cdot) M_{\alpha}^{\dagger} = \sum_{\alpha, \beta=1}^{r} T_{\alpha \gamma}^{\dagger} T_{\alpha \beta} M_{\beta}(\cdot) M_{\alpha}^{\dagger} = \sum_{\beta, \gamma=1}^{r} \delta_{\gamma \beta} M_{\gamma}(\cdot) M_{\beta}^{\dagger} = \sum_{\beta=1}^{r} M_{\beta}(\cdot) M_{\beta}^{\dagger},$$

(S11)

where we have used $\sum_{\alpha=1}^{r} T_{\alpha \gamma}^{\dagger} T_{\alpha \beta} = \delta_{\gamma \beta}$ since $T$ is a unitary matrix.

### III. REPRESENTATIONS OF QUANTUM CHANNELS FOR POVMS

For the measurement operators in Eq. (S6), the matrix representation of the channel in the HS space is

$$\hat{\Phi} = \sum_{\alpha=1}^{r} \sum_{i,j=1}^{d} (c_{\alpha i} |i\rangle \langle i|) \otimes (c_{\alpha j}^{*} |j\rangle \langle j|) = \sum_{i,j=1}^{d} \left( \sum_{\alpha=1}^{r} c_{\alpha i} c_{\alpha j}^{*} \right) |i\rangle \langle j| = \sum_{i,j=1}^{d} c_{i} c_{j}^{*} |i\rangle \langle j|. \quad \text{(S12)}$$

while for the more general case in Eq. (S10), we can similarly obtain

$$\hat{\Phi} = \sum_{\alpha=1}^{r} \sum_{k,l=1}^{s} (\tilde{c}_{\alpha k} \tilde{P}_{k}) \otimes (\tilde{c}_{\alpha l}^{*} \tilde{P}_{l}^{*}) = \sum_{k,l=1}^{s} \tilde{c}_{k}^{*} \tilde{c}_{l} \left( \tilde{P}_{k} \otimes \tilde{P}_{l}^{*} \right) = \sum_{k,l=1}^{s} \tilde{c}_{k}^{*} \tilde{c}_{l} \left( \sum_{i,j \in A_{k}, j \in A_{l}} e^{i(\varphi_{i}-\varphi_{j})} |i\rangle \langle j| \right) \quad \text{(S13)}$$

where $\{\tilde{P}_{k} \otimes \tilde{P}_{l}^{*}\}_{k,l=1}^{s}$ is a set of $s^{2}$ diagonal matrices in HS space that has orthogonal supports, i.e., $(\tilde{P}_{k} \otimes \tilde{P}_{l}^{*})(\tilde{P}_{k'} \otimes \tilde{P}_{l'}^{*}) = \delta_{kk'} \delta_{ll'}(\tilde{P}_{k} \otimes \tilde{P}_{l}^{*})^{2}$. From the Cauchy-Schwarz inequality, $\tilde{c}_{k}^{*} \tilde{c}_{l} < 1$ if $k \neq l$, since $\tilde{c}_{k} \neq e^{i \varphi}$ for any real $\varphi$. So with many applications of the channel,

$$\hat{\Phi}^{m} = \sum_{k,l=1}^{s} (\tilde{c}_{k}^{*} \tilde{c}_{l} m) (\tilde{P}_{k} \otimes \tilde{P}_{l}^{*})^{m} \approx \sum_{k=1}^{s} (\tilde{P}_{k} \otimes \tilde{P}_{k}^{*})^{m} = \sum_{k=1}^{s} \sum_{i,j \in A_{k}} e^{im(\varphi_{i}-\varphi_{j})} |i\rangle \langle j|. \quad \text{(S14)}$$

If $\varphi_{j} = 0$ for any $j \in [1,d]$, i.e., $\tilde{P}_{k} = P_{k}$, then

$$\hat{\Phi}^{m} \approx \sum_{k=1}^{s} P_{k} \otimes P_{k} = \sum_{k=1}^{s} \tilde{P}_{k} = \sum_{k=1}^{s} \sum_{i,j \in A_{k}} |i\rangle \langle j|. \quad \text{(S15)}$$

### IV. TYPICALITY OF SEQUENTIAL POVMS

Since $[M_{\alpha}, M_{\beta}] = 0$ for $\alpha, \beta \in [1,r]$, we can easily prove that $[\hat{M}_{\alpha}, \hat{M}_{\beta}] = 0$. So $\hat{\Phi}^{m}$ can be expanded according to the multinomial theorem,

$$\hat{\Phi}^{m} = \left( \sum_{\alpha=1}^{r} \hat{M}_{\alpha} \right)^{m} = \sum_{\alpha_{1}, \ldots, \alpha_{m}=1}^{r} \hat{M}_{\alpha_{1}} \cdots \hat{M}_{\alpha_{m}} = \sum_{m_{1}, \ldots, m_{r} \geq 0} m! \sum_{m_{1}+\cdots+m_{r}=m} \hat{M}_{1}^{m_{1}} \hat{M}_{2}^{m_{2}} \cdots \hat{M}_{r}^{m_{r}}. \quad \text{(S16)}$$

We define a distribution $F = (f_{1}, \cdots, f_{r}) = (m_{1}/m, \cdots, m_{r}/m)$ to represent the frequencies of each superoperator in $\{\hat{M}_{\alpha}\}_{\alpha=1}^{r}$ to appear in the POVM sequence $\hat{M}_{\alpha_{1}} \cdots \hat{M}_{\alpha_{m}}$, where $\sum_{i=1}^{r} f_{i} = 1$. Then $\hat{\Phi}^{m}$ can be rewritten as

$$\hat{\Phi}^{m} = \sum_{\{F\}} \frac{m!}{(m_{1}f_{1})! \cdots (m_{r}f_{r})!} \hat{M}_{1}^{m_{1}f_{1}} \cdots \hat{M}_{r}^{m_{r}f_{r}}, \quad \text{(S17)}$$
where the summation is over all distributions \{F\} in a \((r - 1)\)-dimensional probability space. Substituting \(\hat{M}_a = \sum_{k,l=1}^{s} c_{atk}^* c_{al}^* (\vec{P}_k \otimes \vec{P}_l^*)\) into Eq. (S17) gives

\[
\hat{\Phi}^m = \sum_{k,l=1}^{s} \sum_{\{F\}} \frac{m!}{(m_{f_1})! \cdots (m_{f_r})!} (c_{1k} c_{1l}^*)^m f_1 \cdots (c_{rk} c_{rl}^*)^m f_r (\vec{P}_k \otimes \vec{P}_r^*)^m
\]

\[
\approx \sum_{k=1}^{s} \sum_{\{F\}} \frac{m!}{(m_{f_1})! \cdots (m_{f_r})!} |c_{1k}|^{2m_{f_1}} \cdots |c_{rk}|^{2m_{f_r}} (\vec{P}_k \otimes \vec{P}_r^*)^m,
\]

(S18)

where we have used \(\langle \vec{P}_k \otimes \vec{P}_l^* \rangle \delta_{kk} \delta_{ll} (\vec{P}_k \otimes \vec{P}_l^*)^2\). To further simplify Eq. (S18), we can use Stirling’s formula \(\ln m! \approx m \ln m - m\) for large \(m\) to obtain

\[
\ln \left( \frac{m!}{(m_{f_1})! \cdots (m_{f_r})!} |c_{1k}|^{2m_{f_1}} \cdots |c_{rk}|^{2m_{f_r}} \right) \approx -m \sum_{i=1}^{r} f_i \ln \frac{f_i}{|c_{ik}|^2} = -mS(F||F_k),
\]

(S19)

where we define \(F_k = (f_{k1}, \cdots, f_{kr}) = (|c_{1k}|^2, \cdots, |c_{rk}|^2)\) with \(c_{1k}, \cdots, c_{rk}\) being entries of \(c_k\) satisfying \(\sum_{i=1}^{r} |c_{ik}|^2 = 1\), and \(S(F||F_k) = \sum_{i=1}^{r} f_i \ln(f_i/f_{ki})\) is the relative entropy between \(F\) and \(F_k\). Then Eq. (S18) is reduced to

\[
\hat{\Phi}^m \approx \sum_{k=1}^{s} \sum_{\{F\}} e^{-mS(F||F_k)} (\vec{P}_k \otimes \vec{P}_k^*)^m.
\]

(S20)

Moreover, for relatively large \(m\), the distribution \(e^{-mS(F||F_k)}\) is concentrated within a small neighborhood around \(F_k\), so \(S(F||F_k) \approx \sum_{i=1}^{r} \frac{(f_i - f_{ki})^2}{2f_{ki}}\), and Eq. (S18) can be further approximated as

\[
\hat{\Phi}^m \approx \sum_{k=1}^{s} \sum_{\{F\}} e^{-\frac{m}{2} \sum_{i=1}^{r} \frac{(f_i - f_{ki})^2}{f_{ki}}} (\vec{P}_k \otimes \vec{P}_k^*)^m.
\]

(S21)

For the special case \(\vec{P}_k = P_k\), Eq. (S20) and Eq. (S22) become

\[
\hat{\Phi}^m \approx \sum_{k=1}^{s} \sum_{\{F\}} e^{-\frac{m}{2} \sum_{i=1}^{r} \frac{(f_i - f_{ki})^2}{f_{ki}}} P_k \approx \sum_{k=1}^{s} \sum_{\{F\}} e^{-\frac{m}{2} \sum_{i=1}^{r} \frac{(f_i - f_{ki})^2}{f_{ki}}} \vec{P}_k,
\]

(S22)

which represents summations of \(s\) Gaussians around \(F_1, \cdots, F_s\), with integration of the \(k\)th Gaussian over the whole probability space giving rise to \(\vec{P}_k\).

For any two Gaussians around \(F_j\) and \(F_k\), they are well separated if the distance between \(F_j\) and \(F_k\) is larger than the sum of the respective Gaussian half widths. In the \((r - 1)\)-dimensional probability space, the straight line connecting \(F_j\) and \(F_k\) is

\[
F_{jk}(t) = (1 - t)F_j + tF_k = ((1 - t)f_{j1} + tf_{k1}, \cdots, (1 - t)f_{jr} + tf_{kr}),
\]

(S23)

where \(t\) is a real number within \([0, 1]\). Define \(\eta\) as the ratio of the minimum hight to the maximum hight within the Gaussian width, then the half widths \(\Delta t_j\) and \(\Delta t_k\) of the two Gaussians along the line \(F_{jk}(t)\) can be derived as

\[
e^{-\frac{m}{2} \sum_{i=1}^{r} \frac{(f_i - f_{ji})^2}{f_{ji}}} = \eta, \quad \Rightarrow \quad \Delta t_j = \sqrt{\frac{2|\ln \eta|}{m}} \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ji}} \right)^{-\frac{1}{2}},
\]

(S24)

\[
e^{-\frac{m}{2} \sum_{i=1}^{r} \frac{(f_i - f_{ki})^2}{f_{ki}}} = \eta, \quad \Rightarrow \quad \Delta t_k = \sqrt{\frac{2|\ln \eta|}{m}} \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ki}} \right)^{-\frac{1}{2}},
\]

(S25)
so the two Gaussians around $F_j$ and $F_k$ are well separated if

$$\Delta t_j + \Delta t_k < 1, \quad \Rightarrow \quad m > 2|\ln \eta| \left[ \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ji}} \right)^{-\frac{1}{2}} + \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ki}} \right)^{-\frac{1}{2}} \right]^2. \quad (S27)$$

For all the Gaussians to be well separated, the lower bound for the number of measurements is

$$m > 2|\ln \eta| \max_{j \neq k, j,k \in [1,s]} \left[ \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ji}} \right)^{-\frac{1}{2}} + \left( \sum_{i=1}^{r} \frac{(f_{ji} - f_{ki})^2}{f_{ki}} \right)^{-\frac{1}{2}} \right]^2. \quad (S28)$$

For the $j$th Gaussian, we define a closed neighborhood $\mathcal{F}_j^\delta$ around $F_j$ in the probability space, such that summation of all the POVM sequences within $\mathcal{F}_j^\delta$ well approximates $\hat{P}_j$. Explicitly, summation of all the POVM sequences within $\mathcal{F}_j^\delta$ gives

$$\hat{\mathcal{P}}_k \approx \sum_{k=1}^{s} \sum_{F \in \mathcal{F}_j^\delta} e^{-mS(F \parallel F_k)} \hat{P}_k = \sum_{k=1}^{s} w_{jk} \hat{P}_k. \quad (S29)$$

where $w_{jk} = \sum_{F \in \mathcal{F}_j^\delta} e^{-mS(F \parallel F_k)}$, satisfying $\sum_{j=1}^{s} w_{jk} \leq 1$ and $w_{jk} \geq 0$. Define $F_j^*$ as a point on the boundary of $\mathcal{F}_j^\delta$ where $F$ takes the minimum on the boundary, then from classical typicality theory [2, 3], we have

$$1 - w_{jj} < \left( \frac{m + r - 1}{m!(r - 1)!} \right) e^{-mS(F \parallel F_j^*)}. \quad (S30)$$

As $\hat{\mathcal{P}}_j^\delta$ and $\hat{P}_j$ are both diagonal operators in the HS space, we can use the trace distance to estimate an upper bound of the error rate in approximating $\hat{P}_j$ with $\hat{\mathcal{P}}_j^\delta$,

$$||\hat{\mathcal{P}}_j^\delta - \hat{P}_j||_1 \approx 1 - w_{jj} + \sum_{k \neq j} w_{jk} \leq \sum_{k=1}^{s} (1 - w_{kk}) < \sum_{k=1}^{s} \left( \frac{m + r - 1}{m!(r - 1)!} \right) e^{-mS(F \parallel F_j^*)}, \quad (S31)$$

which can be made arbitrarily small for a large enough $m$.

[1] S. R. Garcia and R. A. Horn, A Second Course in Linear Algebra (Cambridge University Press, Cambridge, England, 2017).
[2] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley-Interscience, New York, New Jersey, 2006).
[3] J. Mardia, J. Jiao, E. Tzczos, R. D. Nowak, and T. Weissman, Concentration inequalities for the empirical distribution of discrete distributions: beyond the method of types, Inf. Inference 9, 813 (2020).