

On quantum estimation, quantum cloning and finite quantum de Finetti theorems

Giulio Chiribella¹

Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada.

Abstract. This paper presents a series of results on the interplay between quantum estimation, cloning and finite de Finetti theorems. First, we consider the measure-and-prepare channel that uses optimal estimation to convert M copies into k approximate copies of an unknown pure state and we show that this channel is equal to a random loss of all but s particles followed by cloning from s to k copies. When the number k of output copies is large with respect to the number M of input copies the measure-and-prepare channel converges in diamond norm to the optimal universal cloning. In the opposite case, when M is large compared to k , the estimation becomes almost perfect and the measure-and-prepare channel converges in diamond norm to the partial trace over all but k systems. This result is then used to derive de Finetti-type results for quantum states and for symmetric broadcast channels, that is, channels that distribute quantum information to many receivers in a permutationally invariant fashion. Applications of the finite de Finetti theorem for symmetric broadcast channels include the derivation of diamond-norm bounds on the asymptotic convergence of quantum cloning to state estimation and the derivation of bounds on the amount of quantum information that can be jointly decoded by a group of k receivers at the output of a symmetric broadcast channel.

The connection between quantum estimation and cloning is an inspiring leitmotiv of Quantum Information Theory [1–8]. The main related question is: how well can we simulate cloning via estimation? Or, more precisely, how well can we simulate cloning with a “measure-and-prepare” protocol where the input systems are measured, and the output systems are prepared in some state depending on the measurement outcome? As a particular instance of this question, one can ask whether “asymptotic cloning is state estimation” [9], that is, whether the gap between the single-particle fidelity of an optimal cloning channel and the fidelity of the corresponding optimal estimation vanishes when the number of clones tends to infinity.

In Ref. [7] Bae and Acín showed that a channel producing an infinite number of indistinguishable clones must be of the measure-and-prepare

form. On the other hand, Ref. [8] showed that a channel producing a finite number $M < \infty$ of indistinguishable clones can be simulated by a measure-and-prepare channel introducing an error at most of order $\mathcal{O}(1/M)$ on each clone. The proof of Ref. [8] was based on the so-called *finite quantum de Finetti theorem* [10–12], that states that the restriction to k particles of a permutationally invariant M -partite state can be approximated with an error at most of order $\mathcal{O}(k/M)$ by a mixture of product states of the form $\rho^{\otimes k}$. This theorem represents the finite version of the *quantum de Finetti theorem* proved by Caves, Fuchs, and Schack [13] in the context of the Bayesian interpretation of quantum theory. The quantum de Finetti theorem of Ref. [13] corresponds to the ideal $M = \infty$ case and can be directly seen as the quantum formulation of the celebrated de Finetti theorem [14].

Apparently, finite quantum de Finetti theorems are the key to prove the equivalence between asymptotic cloning and estimation. The first result of this paper is to show that, in a sense, the converse is also true: a finite quantum de Finetti-type result can be derived from a particular relation between the optimal estimation [15, 3] and the optimal cloning [2] of an unknown pure state. Precisely, we will see that the optimal measure-and-prepare channel sending M copies of an unknown pure states to k approximate copies is equivalent to a random loss of all but s particles followed by universal cloning from s to k copies. For $M \gg k$ the term with $s = k$ dominates, implying that the optimal measure-and-prepare channel is close to the partial trace over all but k particles. As we will see, this implies directly a de Finetti-type result. Qualitatively, this result shows that the working principle of the finite de Finetti theorems is simply the fact that state estimation from M input copies to k output copies becomes almost perfect when M is large compared to k . Quantitatively, however, the bound derived from the representation of the optimal measure-and-prepare channel as a random mixture of losses followed by cloning can be tightened, as mentioned in subsection 1.4. The bound can be used to derive a finite de Finetti theorem for symmetric quantum broadcast channels, i.e. for channels that distribute quantum information to M indistinguishable users. Examples of symmetric broadcast channels are the channels for the optimal cloning of an unknown state ρ_i randomly drawn with probability p_i from some set of states $\{\rho_i\}$ [16]. The paper concludes with two applications of the finite de Finetti theorem for symmetric broadcast channels. First, the theorem will be used to provide diamond-norm bounds on the asymptotic convergence of quantum cloning to state estimation, thus strengthening the proof of Ref. [8]. As a

second application, the theorem will be used to show that the restriction to k users of any symmetric broadcast channel has a quantum capacity that vanishes at rate $\mathcal{O}(k/M)$ in the large M asymptotics. Even if the overall channel is unitary, and therefore its capacity has the maximum possible value, a group of $k \ll M$ users will only be able to decode a vanishingly small amount of quantum information.

1 The universal measure-and-prepare channel

Let us start with some simple facts about the optimal measure-and-prepare channel transforming M copies of a completely unknown pure states into k approximate copies. The optimal quantum measurement for the estimation of a completely unknown pure state $|\psi\rangle \in \mathcal{H} \simeq \mathbb{C}^d$ from M input copies is given by the *coherent-state POVM* [15, 3]

$$P_{\varphi}^{(M)} \, d\varphi = d_{+}^{(M)} |\varphi\rangle\langle\varphi|^{\otimes M} \, d\varphi \quad d_{+}^{(M)} = \binom{d+M-1}{M} \quad (1)$$

where $|\varphi\rangle \in \mathcal{H}$ is a unit vector and $d\varphi$ is the normalised $SU(d)$ -invariant measure on pure states. This measurement provides a resolution of the identity in the symmetric subspace $(\mathcal{H}^{\otimes M})_{+} \subseteq \mathcal{H}^{\otimes M}$, namely in the subspace spanned by the unit vectors

$$|\mathbf{n}\rangle := \frac{1}{\sqrt{M!n_1!n_2!\dots n_d!}} \sum_{\pi \in S_M} U_{\pi}^{(M)} |1\rangle^{\otimes n_1} |2\rangle^{\otimes n_2} \dots |d\rangle^{\otimes n_d} \quad (2)$$

where $|1\rangle, |2\rangle, \dots, |d\rangle$ is a fixed orthonormal basis for \mathcal{H} , $\mathbf{n} = (n_1, n_2, \dots, n_d)$ is a partition of M , the sum runs over the symmetric group S_M of all permutations of M objects, and $U_{\pi}^{(M)}$ is the unitary operator that permutes the M copies of \mathcal{H} according to the permutation $\pi \in S_M$.

Denoting by $\mathcal{P}_{M,d}$ the set of partitions of M in d nonnegative integers, the normalization of the coherent-state POVM in Eq. (1) is given by

$$\int d\varphi P_{\varphi}^{(M)} = \sum_{\mathbf{n} \in \mathcal{P}_{M,d}} |\mathbf{n}\rangle\langle\mathbf{n}| = P_{+}^{(M)}, \quad (3)$$

where $P_{+}^{(M)}$ is the projector on the symmetric subspace $(\mathcal{H}^{\otimes M})_{+}$.

We now consider the *universal measure-and-prepare channel* from M to k copies, namely the channel that measures the coherent-state POVM $P_{\varphi}^{(M)}$ and, according to the estimate, prepares k copies of the state $|\varphi\rangle$:

$$UMeasPrep_{M,k}(\rho) := \int d\varphi \, \text{Tr}[P_{\varphi}^{(M)} \rho] |\varphi\rangle\langle\varphi|^{\otimes k}. \quad (4)$$

Using Eq. (3) with the substitution $M \rightarrow M+k$ one obtains the equivalent expression

$$\begin{aligned} \mathcal{UMeasPrep}_{M,k}(\rho) &= d_+^{(M)} \int d\varphi \operatorname{Tr}_M \left[\left(\rho \otimes I^{\otimes k} \right) |\varphi\rangle\langle\varphi|^{\otimes M+k} \right] \\ &= \frac{d_+^{(M)}}{d_+^{(M+k)}} \operatorname{Tr}_M \left[\left(\rho \otimes I^{\otimes k} \right) P_+^{(M+k)} \right] \end{aligned} \quad (5)$$

where Tr_M denotes the partial trace over the first M Hilbert spaces.

For an arbitrary pure state $|\psi\rangle$, the fidelity between the channel output $\mathcal{UMeasPrep}_{M,k}(|\psi\rangle\langle\psi|^{\otimes M})$ and the desideratum $|\psi\rangle\langle\psi|^{\otimes k}$ is given by $F_{M,k} = d_+^{(M)}/d_+^{(M+k)}$, as it is immediate from Eq. (5). In fact, it is easy to show that $F_{M,k} = d_+^{(M)}/d_+^{(M+k)}$ is the maximum average fidelity achievable with a measure-and-prepare channel $\mathcal{M}(\rho) = \sum_i \operatorname{Tr}[P_i\rho]\rho_i$, where $\{P_i\}$ is a POVM on $(\mathcal{H}^{\otimes M})_+$ and $\{\rho_i\}$ is a set of states on $(\mathcal{H}^{\otimes k})_+$. Indeed, in this case one has

$$\begin{aligned} \bar{F} &= \int d\psi \langle\psi|^{\otimes k} \mathcal{M}(|\psi\rangle\langle\psi|^{\otimes M}) |\psi\rangle^{\otimes k} = \frac{\sum_i \operatorname{Tr} \left[(P_i \otimes \rho_i) P_+^{(M+k)} \right]}{d_+^{(M+k)}} \\ &\leq \frac{\sum_i \operatorname{Tr} [P_i \otimes \rho_i]}{d_+^{(M+k)}} = \frac{d_+^{(M)}}{d_+^{(M+k)}} \end{aligned}$$

(cf. Bruß and Macchiavello [3] for the $k = 1$ case). Clearly, when M is large compared to k the fidelity $F_{M,k}$ is close to unit: the desired output states $|\psi\rangle^{\otimes k}$ are much less distinguishable than the input states $|\psi\rangle^{\otimes M}$, thus allowing for an almost ideal re-preparation. In this case, one has

$$\mathcal{UMeasPrep}_{M,k}(|\psi\rangle\langle\psi|^{\otimes M}) \approx |\psi\rangle\langle\psi|^{\otimes k} \quad \forall |\psi\rangle \in \mathcal{H},$$

or, equivalently (cf. the Appendix),

$$\mathcal{UMeasPrep}_{M,k}(\rho) \approx \operatorname{Tr}_{M-k}[\rho] \quad \forall \rho \in \operatorname{Lin} \left((\mathcal{H}^{\otimes M})_+ \right),$$

where $\operatorname{Lin}(V)$ denotes the set of linear operators on the linear space V ($V = (\mathcal{H}^{\otimes M})_+$ in this case). Despite the simplicity of the above observation, the consequences of the fact that for $M \gg k$ the estimation from M to k copies is “almost ideal” are far from trivial: as we will see, this simple fact can be considered as the working principle of the finite de Finetti theorems.

The purpose of the next subsection is to give a convenient representation of the channel $\mathcal{UMeasPrep}_{M,k}$ as a convex mixture of losses concatenated with cloning channels. Using this representation we will show that in the limit $k/M \rightarrow 0$ the channel $\mathcal{UMeasPrep}_{M,k}$ converges to the partial trace Tr_{M-k} in the strongest possible sense, in terms of the *diamond norm* [17], equivalent to the *norm of complete boundedness* [18] of the channel in Heisenberg picture. Operationally, convergence in the diamond norm means that for $M \gg k$ the two channels $\mathcal{UMeasPrep}_{M,k}$ and Tr_{M-k} are almost indistinguishable even when entanglement-assisted discrimination strategies are employed.

1.1 Representation of the universal measure-and-prepare channel as a mixture of universal cloning channels

The main result of this subsection is the following expression, proved in the Appendix:

$$\mathcal{UMeasPrep}_{M,k}(\rho) = \sum_{s=0}^{\min\{k,M\}} p_s \mathcal{UClon}_{s,k}(\text{Tr}_{M-s}[\rho]), \quad p_s = \frac{\binom{M}{s} \binom{d+k-1}{k-s}}{\binom{d+M+k-1}{k}}, \quad (6)$$

$\mathcal{UClon}_{s,k}$ being the *universal s-to-k cloning channel*, i.e. the optimal quantum channel that clones an unknown pure state $|\psi\rangle$ from s to k copies, given by [2, 4]

$$\mathcal{UClon}_{s,k}(\rho) = \frac{d_+^{(s)}}{d_+^{(k)}} P_+^{(k)} \left(\rho \otimes I^{\otimes(k-s)} \right) P_+^{(k)}. \quad (7)$$

Note that $\{p_s\}$ is a probability distribution, as the normalization

$$\sum_{s=0}^{\min\{k,M\}} p_s = \sum_{s=0}^k p_s = 1$$

follows immediately from the fact that $p_s = 0$ if $s > M$ and from the Chu-Vandermonde convolution formula (see Eq. (7.6) p. 59 of Ref. [19] for an equivalent formula)

$$\binom{z+w}{N} = \sum_{i=0}^N \binom{z}{i} \binom{w}{N-i} \quad \forall z, w \in \mathbb{C}, \forall N \in \mathbb{N}. \quad (8)$$

Eq. (6) means that measuring M copies and re-preparing k copies has the same effect of a random loss of $M - s$ systems followed by quantum cloning from s to k copies: the particles that are missing are replaced by clones.

In the following we will consider the two extreme cases $k \gg M$ and $M \gg k$. In the former, we will see that the measure-and-prepare channel $\mathcal{U}MeasPrep_{M,k}$ converges to the universal cloning $\mathcal{U}Clon_{M,k}$. In the latter, the measure-and-prepare channel $\mathcal{U}MeasPrep_{M,k}$ will converge to the partial trace Tr_{M-k} , leading to a de Finetti-type result. The convergence will be quantified in terms of the *diamond norm* [17] (in Heisenberg picture, the completely bounded norm [18]), which for a Hermitian-preserving map Δ from $\text{Lin}(\mathcal{H}_{in})$ to $\text{Lin}(\mathcal{H}_{out})$ is given by

$$\|\Delta\|_{\diamond} = \sup_{\mathcal{H}_A} \sup_{|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_{in}, \|\Psi\|=1} \|(\mathcal{I}_A \otimes \Delta)(|\Psi\rangle\langle\Psi|)\|_1, \quad (9)$$

where $\|A\|_1 = \text{Tr}|A|$ is the trace-norm and \mathcal{I}_A is the identity map on the ancillary Hilbert space \mathcal{H}_A .

1.2 $k \gg M$ case: convergence to universal cloning

Suppose that the number of output copies k is larger than the number of input copies M . In the limit of $M/k \rightarrow 0$, the term with $s = M$ in Eq. (6) dominates, thus giving $\mathcal{U}MeasPrep_{M,k} \approx \mathcal{U}Clon_{M,k}$.

An estimate of the diamond-norm convergence to universal cloning is given by the following:

Theorem 1 (Convergence to universal cloning). *The universal measure-and-prepare channel $\mathcal{U}MeasPrep_{M,k}$ converges to the universal cloning channel $\mathcal{U}Clon_{M,k}$ in the limit $k \rightarrow \infty$. In particular, the following bound holds:*

$$\|\mathcal{U}MeasPrep_{M,k} - \mathcal{U}Clon_{M,k}\|_{\diamond} \leq \frac{2M(d + M - 1)}{k + d}. \quad (10)$$

Proof. Writing $\mathcal{U}MeasPrep_{M,k} = p_M \mathcal{U}Clon_{M,k} + (1 - p_M) \mathcal{R}est$ where $\mathcal{R}est$ is a suitable channel, one has $\|\mathcal{U}MeasPrep_{M,k} - \mathcal{U}Clon_{M,k}\|_{\diamond} \leq (1 - p_M) \|\mathcal{R}est - \mathcal{U}Clon_{M,k}\|_{\diamond}$. Since the distance between the two channels $\mathcal{R}est$ and $\mathcal{U}Clon_{M,k}$ is upper bounded by 2, this gives $\|\mathcal{U}MeasPrep_{M,k} - \mathcal{U}Clon_{M,k}\|_{\diamond} \leq 2(1 - p_M)$. The bound in Eq. (11) just comes from a lower

bound on p_M :

$$\begin{aligned} p_M &= \frac{k(k-1)\dots(k-M+1)}{(d+M+k-1)(d+M+k-2)\dots(d+k)} \geq \left(\frac{k-M+1}{d+k}\right)^M \\ &= \left(1 - \frac{d+M-1}{d+k}\right)^M \geq 1 - \frac{M(d+M-1)}{d+k}. \end{aligned}$$

□

Theorem 1 shows an exceptionally strong case of equivalence between asymptotic cloning and state estimation: it shows that, in the universal case, the optimal cloning channel [2, 4] converges in diamond norm to the measure-and-prepare channel $\mathcal{UMeasPrep}_{M,k}$ when the number k of output copies is large with respect to the number M of input copies. It is worth stressing, however, that this result is very specific to the universal case. What can be proved for generic (i.e. non-universal) cloning channels is that the k -particle restrictions of a cloning channel with M output copies can be simulated by a measure-and-prepare channel with an error of order k/M (see subsection 2.2). This result will emerge from the analysis of Eq. (6) in the $M \gg k$ case, which is discussed in the next subsection.

1.3 $M \gg k$ case: convergence to the partial trace

Here we consider the case where the number of input copies k is large with respect to the number of output copies M . In this case, the leading term in Eq. (6) is the term with $s = k$. Note that, since for $s = k$ the universal cloning $\mathcal{UClon}_{k,k}$ is simply the identity map on $(\mathcal{H}^{\otimes k})_+$, the corresponding term in Eq. (6) is the partial trace Tr_{M-k} . Therefore, when M is large compared to k the channel $\mathcal{UMeasPrep}_{M,k}$ converges to the trace Tr_{M-k} . This implies an almost ideal estimation, with $\mathcal{UMeasPrep}_{M,k}(|\psi\rangle\langle\psi|^{\otimes M}) \approx \text{Tr}_{M-k}[|\psi\rangle\langle\psi|^{\otimes M}] = |\psi\rangle\langle\psi|^{\otimes k}$. A first estimate on the diamond-norm convergence to ideal estimation is given by the following

Theorem 2 (Convergence to ideal estimation). *The universal measure-and-prepare channel $\mathcal{UMeasPrep}_{M,k}$ converges to the trace channel Tr_{M-k} in the limit $M \rightarrow \infty$. In particular, the following bound holds*

$$\|\mathcal{UMeasPrep}_{M,k} - \text{Tr}_{M-k}\|_{\diamond} \leq \frac{2k(d+k-1)}{M+d}. \quad (11)$$

Proof. Writing $\mathcal{UMeasPrep}_{M,k} = p_k \text{Tr}_{M-k} + (1-p_k) \mathcal{Rest}$ where \mathcal{Rest} is a suitable channel, one has $\|\mathcal{UMeasPrep}_{M,k} - \text{Tr}_{M-k}\|_{\diamond} \leq (1-p_k) \|\mathcal{Rest} -$

$\text{Tr}_{M-k}\|\diamond$. Since the distance between the two channels $\mathcal{R}est$ and Tr_{M-k} is upper bounded by 2, this gives $\|\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k}\|\diamond \leq 2(1 - p_k)$. The bound in Eq. (11) just comes from a lower bound on p_k :

$$\begin{aligned} p_k &= \frac{M(M-1)\dots(M-k+1)}{(d+M+k-1)(d+M+k-2)\dots(d+M)} \geq \left(\frac{M-k+1}{d+M}\right)^k \\ &= \left(1 - \frac{d+k-1}{d+M}\right)^k \geq 1 - \frac{k(d+k-1)}{d+M}. \end{aligned}$$

□

The bound of Eq. (11) clearly implies a de Finetti-type result:

Corollary 1. *For every state ρ with support in the symmetric space $(\mathcal{H}^{\otimes M})_+$ there exists a state $\tilde{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|^{\otimes M}$ such that the k -particle restrictions of ρ and $\tilde{\rho}$ are almost indistinguishable for large M . Precisely, denoting the k -particle restrictions by $\rho^{(k)} = \text{Tr}_{M-k}[\rho]$ and $\tilde{\rho}^{(k)} = \text{Tr}_{M-k}[\tilde{\rho}]$, one has*

$$\left\|\rho^{(k)} - \tilde{\rho}^{(k)}\right\|_1 \leq \frac{2k(d+k-1)}{M+d} \quad (12)$$

Proof. Taking $\tilde{\rho} = \mathcal{U}MeasPrep_{M,M}(\rho)$ we obtain a state of the desired form, and, in addition, we have

$$\begin{aligned} \left\|\tilde{\rho}^{(k)} - \rho^{(k)}\right\|_1 &= \left\|\mathcal{U}MeasPrep_{M,k}(\rho) - \text{Tr}_{M-k}[\rho]\right\|_1 \\ &\leq \left\|\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k}\right\|_\diamond \\ &\leq \frac{2k(d+k-1)}{M+d}. \end{aligned}$$

□

The bound of Eq. (12) can be extended to the case of states on $\mathcal{H}^{\otimes M}$ that are just permutationally invariant, using the fact that *i*) every permutationally invariant state on $\mathcal{H}^{\otimes M}$ has a purification in the symmetric space $(\mathcal{K}^{\otimes M})_+$, with $\mathcal{K} = \mathcal{H} \otimes \mathcal{H}$ (see e.g. [10]) and that *ii*) the norm is non-increasing under partial traces. Therefore, for a permutationally invariant state the bound of Eq. (12) holds with the substitution $d \rightarrow d^2$.

1.4 Improving the bound

The bound of Eq. (11) provides good estimates for $k = 1$ or when d is large, so that $Mk \leq d^2$ (see the observation below). Outside this range of values, the estimate can be improved using the technique developed in Ref. [10] for the proof of the finite de Finetti theorem, combined with the bounding of Ref. [8]:

Theorem 3. *The universal measure-and-prepare channel $\mathcal{U}MeasPrep_{M,k}$ satisfies the bound*

$$\|\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k}\|_{\diamond} \leq 4 \left(1 - \sqrt{\frac{d_+^{(M-k)}}{d_+^{(M)}}} \right) \leq \frac{2kd}{M} \quad (13)$$

Observation. Note that the quantity $2kd/M$ in Eq. (13) is larger than the quantity $2k(d+k-1)/(M+d)$ in Eq. (11) whenever $M(k-1) \leq d^2$. In general, the more accurate estimate is obtained by taking the minimum between the two quantities in Eqs. (11) and (13).

Proof of Theorem 2. Let $|\Psi\rangle$ be an arbitrary state in $\mathcal{H}_A \otimes (\mathcal{H}^{\otimes M})_+$, where \mathcal{H}_A is an arbitrary Hilbert space. Define the states

$$\begin{aligned} \rho^{(Ak)} &= (\mathcal{I}_A \otimes \text{Tr}_{M-k}) [|\Psi\rangle\langle\Psi|] \\ \tilde{\rho}^{(Ak)} &= (\mathcal{I}_A \otimes \mathcal{U}MeasPrep_{M,k}) [|\Psi\rangle\langle\Psi|]. \end{aligned}$$

Using the normalization of the coherent-state POVM in Eq. (3) with the substitution $M \rightarrow M-k$, we can write $\rho^{(Ak)} = \int d\varphi \rho_{\varphi}^{(Ak)}$, where

$$\rho_{\varphi}^{(Ak)} = \text{Tr}_{M-k} \left[|\Psi\rangle\langle\Psi| \left(I_A \otimes I^{\otimes k} \otimes P_{\varphi}^{(M-k)} \right) \right].$$

On the other hand, the state $\tilde{\rho}^{(Ak)}$ can be written as

$$\tilde{\rho}^{(Ak)} = \lambda \int d\varphi \left(I_A \otimes P_{\varphi}^{(k)} \right) \rho_{\varphi}^{(Ak)} \left(I_A \otimes P_{\varphi}^{(k)} \right),$$

with $\lambda = \frac{d_+^{(M)}}{d_+^{(M-k)} d_+^{(k)2}}$. The difference between $\rho^{(Ak)} - \tilde{\rho}^{(Ak)}$ is then given by

$$\rho^{(Ak)} - \tilde{\rho}^{(Ak)} = \int d\varphi (A_{\varphi} - B_{\varphi} A_{\varphi} B_{\varphi}),$$

where $A_{\varphi} = \rho_{\varphi}^{(Ak)}$ and $B_{\varphi} = \sqrt{\lambda} \left(I_A \otimes P_{\varphi}^{(k)} \right)$.

Using the relation $A - BAB = A(I-B) + (I-B)A - (I-B)A(I-B)$ we obtain

$$\rho^{(Ak)} - \tilde{\rho}^{(Ak)} = C + C^{\dagger} - D, \quad (14)$$

where $C = \int d\varphi A_{\varphi} (I - B_{\varphi})$ and $D = \int d\varphi (I - B_{\varphi}) A_{\varphi} (I - B_{\varphi})$. The operator C can be calculated using the relation

$$\begin{aligned} \int d\varphi A_{\varphi} B_{\varphi} &= \frac{\sqrt{\lambda} d_+^{(k)} d_+^{(M-k)}}{d_+^{(M)}} \int d\varphi \text{Tr}_{M-k} \left[|\Psi\rangle\langle\Psi| \left(I_A \otimes P_{\varphi}^{(M)} \right) \right] \\ &= \sqrt{\frac{d_+^{(M-k)}}{d_+^{(M)}}} \text{Tr}_{M-k} [|\Psi\rangle\langle\Psi|] = \sqrt{\frac{d_+^{(M-k)}}{d_+^{(M)}}} \rho^{(Ak)}, \end{aligned}$$

which gives $C = \left(1 - \sqrt{d_+^{(M-k)}/d_+^{(M)}}\right) \rho^{(Ak)} = C^\dagger$.

Taking the norm on both sides of Eq. (14), using the triangle inequality, and the fact that C and D are both nonnegative we obtain $\|\rho^{(Ak)} - \tilde{\rho}^{(Ak)}\|_1 \leq 2\|C\|_1 + \|D\|_1 = 2\text{Tr}[C] + \text{Tr}[D]$. Finally, taking the trace on both sides of Eq. (14) we get $\text{Tr}[D] = 2\text{Tr}[C]$. The inequality $\|\rho^{(Ak)} - \tilde{\rho}^{(Ak)}\|_1 \leq 4\text{Tr}[C]$ then gives the first bound in Eq. (13). The second bound follows from the inequalities $d_+^{(M-k)}/d_+^{(M)} \geq (1 - k/M)^d$ (see e.g. Ref.[10]) and $(1 - x)^\alpha \geq 1 - \alpha x$, which holds for $\alpha \geq 1$ and $x \leq 1$. \square

2 Symmetric broadcast channels

A *quantum broadcast channel* is a channel with a single sender and many receivers [20]. We define a *symmetric* broadcast channel as a channel where the Hilbert spaces of all receivers are isomorphic and the output of the channel is invariant under permutations. Precisely, we say that a channel $\mathcal{E} : \text{Lin}(\mathcal{H}_{in}) \rightarrow \text{Lin}(\mathcal{H}^{\otimes M})$ is a *symmetric broadcast channel* if

$$\mathcal{E} = \mathcal{U}_\pi^{(M)} \mathcal{E} \quad \forall \pi \in S_M, \quad (15)$$

where $\mathcal{U}_\pi^{(M)}$ is the unitary channel defined by $\mathcal{U}_\pi^{(M)}(\rho) := U_\pi^{(M)} \rho U_\pi^{(M)\dagger}$, $\rho \in \text{Lin}(\mathcal{H}_{in})$. The requirement of Eq. (15) models the situation where the quantum information in the input is equally spread over all receivers: any possible permutation of the receivers leaves the channel invariant. An example of symmetric broadcast channel is the optimal cloning channel for an arbitrary set of pure states, whenever the figure of merit is the average of the single-copy fidelity over all the M output copies (see e.g. [4]). In the following we will prove a finite de Finetti theorem for symmetric broadcast channels. The theorem is then used to show a strong form of the equivalence between asymptotic cloning and state estimation and to provide bounds on the amount of quantum information that can be jointly decoded by k receivers at the output of a symmetric broadcast channel.

2.1 Finite de Finetti theorems for symmetric quantum broadcast channels

For symmetric broadcast channels with output in the symmetric subspace the following approximation result holds:

Theorem 4 (Finite de Finetti theorem for symmetric broadcast channels with output in the symmetric subspace). *For a symmet-*

ric broadcast channel $\mathcal{E} : \text{Lin}(\mathcal{H}_{in}) \rightarrow \text{Lin}((\mathcal{H}^{\otimes M})_+)$ there is a measure-and-prepare channel $\tilde{\mathcal{E}}$ of the form $\tilde{\mathcal{E}}(\rho) = \sum_i \text{Tr}[P_i \rho] |\psi_i\rangle\langle\psi_i|^{\otimes M}$ such that

$$\|\tilde{\mathcal{E}}^{(k)} - \mathcal{E}^{(k)}\|_{\diamond} \leq 4 \left(1 - \sqrt{\frac{d_+^{(M-k)}}{d_+^{(M)}}} \right) \leq \frac{2kd}{M}, \quad (16)$$

where $\tilde{\mathcal{E}}^{(k)} := \text{Tr}_{M-k} \circ \tilde{\mathcal{E}}$ and $\mathcal{E}^{(k)} := \text{Tr}_{M-k} \circ \mathcal{E}$.

Proof Define the measure-and-prepare channel $\tilde{\mathcal{E}}$ as

$$\tilde{\mathcal{E}}(\rho) = \mathcal{U} \text{MeasPrep}_{M,M} \circ \mathcal{E}(\rho) = \int d\varphi \text{Tr}[Q_{\varphi} \rho] |\varphi\rangle\langle\varphi|^{\otimes M},$$

where $Q_{\varphi} d\varphi$ is the POVM defined by

$$\text{Tr}[Q_{\varphi} \rho] = \text{Tr}[P_{\varphi}^{(M)} \mathcal{E}(\rho)] \quad \forall \rho \in \text{Lin}(\mathcal{H}_{in}),$$

that is, $Q_{\varphi} d\varphi$ is the POVM obtained by applying the channel \mathcal{E} in Heisenberg picture to the coherent-state POVM $P_{\varphi}^{(M)} d\varphi$. From the definition of $\tilde{\mathcal{E}}$ it is clear that $\mathcal{E}^{(k)} = \mathcal{U} \text{MeasPrep}_{M,k} \circ \mathcal{E}$. Using the submultiplicativity property $\|\mathcal{A}\mathcal{B}\|_{\diamond} \leq \|\mathcal{A}\|_{\diamond} \|\mathcal{B}\|_{\diamond}$, the fact that $\|\mathcal{E}\|_{\diamond} = 1$ since \mathcal{E} is a channel, and the bound of Eq. (13) we then obtain

$$\begin{aligned} \|\tilde{\mathcal{E}}^{(k)} - \mathcal{E}^{(k)}\|_{\diamond} &= \|(\mathcal{U} \text{MeasPrep}_{M,k} - \text{Tr}_{M-k}) \circ \mathcal{E}\|_{\diamond} \\ &\leq 4 \left(1 - \sqrt{\frac{d_+^{(M)}}{d_+^{(M+k)}}} \right) \leq \frac{2dk}{M}. \end{aligned}$$

□

The extension to arbitrary broadcast channels with permutationally invariant output is given in the following

Theorem 5 (Finite de Finetti theorem for symmetric broadcast channels). *For every symmetric broadcast channel $\mathcal{E} : \text{Lin}(\mathcal{H}_{in}) \rightarrow \text{Lin}(\mathcal{H}^{\otimes M})$ there is a measure-and-prepare channel $\tilde{\mathcal{E}} = \sum_i \text{Tr}[P_i \rho] \rho_i^{\otimes M}$ such that the bounds in Eq. (16) hold with the substitution $d \rightarrow d^2$.*

Proof Consider the Stinespring dilation $\mathcal{E}(\rho) = \text{Tr}_{env}[V \rho V^{\dagger}]$, where $V : \mathcal{H}_{in} \rightarrow \mathcal{H}^{\otimes M} \otimes \mathcal{H}_{env}$ is an isometry and Tr_{env} is the partial trace over the environment Hilbert space \mathcal{H}_{env} . Since by definition a symmetric broadcast channel satisfies the relation

$$\mathcal{E}(\rho) = U_{\pi}^{(M)} \mathcal{E}(\rho) U_{\pi}^{(M)}, \quad \forall \rho \in \text{Lin}(\mathcal{H}_{in}), \forall \pi \in S_M,$$

it follows from the theory of covariant channels that one can choose $\mathcal{H}_{env} = \mathcal{H}^{\otimes M} \otimes \mathcal{H}_{in}$ and V with the property

$$\left(U_{\pi}^{(M)} \otimes U_{\pi}^{(M)} \otimes I_{in} \right) V = V, \quad \forall \pi \in S_M$$

(see Eq. (65) of Ref. [21]). This property implies that the output of the isometric channel $\mathcal{V}(\rho) = V\rho V^\dagger$ has support in the subspace $(\mathcal{K}^{\otimes M})_+ \otimes \mathcal{H}_{in}$, where $\mathcal{K} = \mathcal{H}^{\otimes 2}$. Now, consider the channel $\mathcal{F} = \text{Tr}_{in} \circ \mathcal{V} : \text{Lin}(\mathcal{H}_{in}) \rightarrow \text{Lin}((\mathcal{K}^{\otimes M})_+)$. By theorem 4, there exists a measure-and-prepare channel $\tilde{\mathcal{F}}$ of the form $\tilde{\mathcal{F}}(\rho) = \sum_i \text{Tr}[P_i \rho] |\Psi_i\rangle\langle\Psi_i|^{\otimes M}$, with $|\Psi_i\rangle \in \mathcal{H}^{\otimes 2}$, such that the restrictions $\mathcal{F}^{(k)}$ and $\tilde{\mathcal{F}}^{(k)}$ satisfy the bound of Eq. (16) with the substitution $d \rightarrow d^2$. To obtain the desired result it is sufficient to define the channel $\tilde{\mathcal{E}}$ as $\tilde{\mathcal{E}}(\rho) = \text{Tr}_{env}[\tilde{\mathcal{V}}(\rho)] = \sum_i \text{Tr}[P_i \rho] \rho_i^{\otimes M}$, where ρ_i is the reduced density matrix of $|\Psi_i\rangle\langle\Psi_i|$, and to use the relation

$$\begin{aligned} \|\tilde{\mathcal{E}}^{(k)} - \mathcal{E}^{(k)}\|_{\diamond} &= \|\text{Tr}_{env,k} \circ (\tilde{\mathcal{F}}^{(k)} - \mathcal{F}^{(k)})\|_{\diamond} \\ &\leq \|\tilde{\mathcal{F}}^{(k)} - \mathcal{F}^{(k)}\|_{\diamond}, \end{aligned}$$

where $\text{Tr}_{env,k}$ denotes the partial trace over the k systems in the environment. \square

Observation. The usual de Finetti theorems for quantum states [10–12] can be retrieved from theorems 4 and 5 in the special case of symmetric broadcasting channels with trivial input space $\mathcal{H}_{in} \simeq \mathbb{C}$. In this case the POVM $\{P_i\}$ becomes just a collection of probabilities $\{p_i\}$.

Theorems 4 and 5 have many interesting consequences: first of all they imply that the output state of k receivers contains a vanishing amount of entanglement in the limit of vanishing k/M . Moreover, they imply that the information transmitted to a small number of receivers can only be classical, while the amount of quantum information is vanishing. This observation will be made quantitatively precise in subsection 2.3. Another consequence is a strong form of the equivalence between asymptotic cloning states estimation, briefly discussed in the next subsection.

2.2 Strong equivalence between asymptotic pure state cloning and state estimation

Let $\{|\psi_x\rangle\}_{x \in X} \subset \mathcal{H}$ be a set of pure states and $\{p_x\}$ a corresponding set of prior probabilities. An N -to- M cloning channel transforms N copies of a state $|\psi_x\rangle$ into M approximate copies, the joint state of the copies being a state on $\mathcal{H}^{\otimes M}$. The requirement that each single copy have the same

fidelity with the state $|\psi_x\rangle$ is implemented without loss of generality by taking cloning channels with permutationally invariant output: clearly, such cloning channels are an example of symmetric broadcast channels. Let us call $\mathcal{C}lon_{N,M}$ the N -to- M cloning channel under consideration and let $\widetilde{\mathcal{C}lon}_{N,M}$ be the measure-and-prepare channel defined in Theorem 5. Theorem 5 then implies the bound

$$\left\| \mathcal{C}lon_{N,M}^{(k)} - \widetilde{\mathcal{C}lon}_{N,M}^{(k)} \right\|_{\diamond} \leq \frac{2d^2k}{M}, \quad (17)$$

that is, for fixed k and d the cloning channel becomes more and more indistinguishable from a measure-and-prepare channel as M increases. In particular, if $\mathcal{C}lon_{N,M}$ is the optimal cloning channel according to some figure of merit, Eq. (17) entails the convergence of optimal cloning to estimation. Note that the convergence in diamond norm represents an improvement over the trace-norm convergence of Ref. [8], as it states that cloning is indistinguishable from estimation even with the aid of entanglement with a reference system. The convergence of the fidelities is then a simple corollary: For every state ψ_x , the single-copy fidelity is given by

$$F_{clon}[N, M, x] = \langle \psi_x | \mathcal{C}lon_{N,M}^{(1)}(|\psi_x\rangle\langle\psi_x|^{\otimes M}) | \psi_x \rangle.$$

Denoting by $F_{\widetilde{clon}}[N, x]$ the single-copy fidelity for the measure-and-prepare channel $\widetilde{\mathcal{C}lon}_{N,M}$ (note that in this case the fidelity is independent of M), we have

$$\begin{aligned} |F_{clon}[N, M, x] - F_{\widetilde{clon}}[N, x]| &\leq \left\| (\mathcal{C}lon_{N,M}^{(1)} - \widetilde{\mathcal{C}lon}_{N,M}^{(1)}) (|\psi_x\rangle\langle\psi_x|^{\otimes N}) \right\|_1 \\ &\leq \left\| \mathcal{C}lon_{N,M}^{(1)} - \widetilde{\mathcal{C}lon}_{N,M}^{(1)} \right\|_{\diamond} \leq \frac{2d^2k}{M}. \end{aligned}$$

Denoting by $F_{est}[N]$ the maximum average fidelity achievable by a measure-and-prepare channel and using the fact that $F_{est}[N] \leq F_{clon}[N, M], \forall M$ we then have the bound

$$\begin{aligned} 0 \leq F_{clon}[N, M] - F_{est}[N] &\leq \left| \sum_x p_x (F_{clon}[N, M, x] - F_{\widetilde{clon}}[N, x]) \right| \\ &\leq \sum_x p_x |F_{clon}[N, M, x] - F_{\widetilde{clon}}[N, x]| \leq \frac{2d^2k}{M}, \end{aligned}$$

which implies the limit $\lim_{M \rightarrow \infty} F_{clon}[N, M] = F_{est}[N]$.

2.3 Bounds on the quantum capacities of the k -receivers restriction of a symmetric broadcast channel

Theorems 4 and 5 also imply a set of bounds on the amount of quantum information that k receivers can jointly decode at the output of a symmetric broadcast channel \mathcal{E} . For definiteness, let us consider the case of a channel \mathcal{E} with output in the symmetric subspace $(\mathcal{H}^{\otimes M})_+$: this is the case, e.g. of all known examples of optimal pure state cloning [16]. A first bound on the quantum capacity comes from the continuity result of Ref.[22], that, along with the fact that measure-and-prepare channels have zero quantum capacity, yields the following estimate

$$Q(\mathcal{E}^{(k)}) = |Q(\mathcal{E}^{(k)}) - Q(\tilde{\mathcal{E}}^{(k)})| \leq \frac{16kd}{M} \log d_+^{(k)} + 4H\left(\frac{2kd}{M}\right). \quad (18)$$

where H is the binary entropy $H(x) = -x \log x - (1-x) \log(1-x)$, and \log denotes the logarithm in base 2.

Two other estimates are given in the following

Corollary 2. *The quantum capacity of the k -receivers restriction of a symmetric broadcast channel $\mathcal{E} : \text{Lin}(\mathcal{H}_{in}) \rightarrow \text{Lin}((\mathcal{H}^{\otimes M})_+)$ satisfies the bound*

$$Q(\mathcal{E}^{(k)}) \leq \min \left\{ \log \left(1 + \frac{2kdd_+^{(k)}}{M} \right), \log \left(1 + \frac{2kdd_{in}}{M} \right) \right\} \quad (19)$$

$$\leq \min \left\{ \frac{2kdd_+^{(k)}}{M}, \frac{2kdd_{in}}{M} \right\} \quad (20)$$

Proof Holevo and Werner proved that the quantum capacity of a channel \mathcal{C} is upper bounded by the ε -quantum capacity $Q_\varepsilon(\mathcal{C})$ [23] (i.e. the supremum of the rates that are asymptotically achievable with error bounded by ε), and that $Q_\varepsilon(\mathcal{C})$ is upper bounded by $\log \|\mathcal{C}\Theta_{in}\|_\diamond$, where Θ_{in} is the transposition map on the input space \mathcal{H}_{in} . We then obtain

$$\begin{aligned} Q(\mathcal{E}^{(k)}) &\leq Q_\varepsilon(\mathcal{E}^{(k)}) \leq \log \|\tilde{\mathcal{E}}^{(k)}\Theta_{in} + (\mathcal{E}^{(k)} - \tilde{\mathcal{E}}^{(k)})\Theta_{in}\|_\diamond \\ &\leq \log \left(\|\tilde{\mathcal{E}}^{(k)}\Theta_{in}\|_\diamond + \|\mathcal{E}^{(k)} - \tilde{\mathcal{E}}^{(k)}\|_\diamond \|\Theta_{in}\|_\diamond \right) \\ &\leq \log \left(1 + \frac{2kdd_{in}}{M} \right), \end{aligned}$$

having used the triangle inequality, the submultiplicativity $\|\mathcal{A}\mathcal{B}\|_\diamond \leq \|\mathcal{A}\|_\diamond \|\mathcal{B}\|_\diamond$ the fact that $\|\tilde{\mathcal{E}}^{(k)}\Theta_{in}\|_\diamond = 1$ since $\tilde{\mathcal{E}}^{(k)}\Theta_{in}(\rho) = \int d\varphi \text{Tr}[Q_\varphi^T \rho] |\varphi\rangle\langle\varphi|^{\otimes k}$ is

still a quantum channel, the equality $\|\Theta_{in}\|_{\diamond} = d_{in}$, and the bound of Eq. (16). Similarly, denoting by $\Theta_+^{(M)}$ and $\Theta_+^{(k)}$ the transposition maps on $(\mathcal{H}^{\otimes M})_+$ and $(\mathcal{H}^{\otimes k})_+$, respectively, we obtain

$$\begin{aligned}
Q(\mathcal{E}^{(k)}) &\leq Q_{\epsilon}(\mathcal{E}^{(k)}) \leq \log \|\tilde{\mathcal{E}}^{(k)}\Theta_+ + (\mathcal{E}^{(k)} - \tilde{\mathcal{E}}^{(k)})\Theta_{in}\|_{\diamond} \\
&\leq \log \left[1 + \|(\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k})\Theta_+^{(M)}(\Theta_+^{(M)}\mathcal{E}\Theta_{in})\|_{\diamond} \right] \\
&\leq \log \left[1 + \|(\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k})\Theta_+^{(M)}\|_{\diamond} \right] \\
&\leq \log \left[1 + \|\Theta_+^{(k)}\|_{\diamond} \|\Theta_+^{(k)}(\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k})\Theta_+^{(M)}\|_{\diamond} \right] \\
&= \log \left[1 + \|\Theta_+^{(k)}\|_{\diamond} \|\mathcal{U}MeasPrep_{M,k} - \text{Tr}_{M-k}\|_{\diamond} \right] \\
&\leq \log \left(1 + \frac{2kdd_+^{(k)}}{M} \right).
\end{aligned}$$

having used the triangle inequality, the submultiplicativity $\|\mathcal{A}\mathcal{B}\|_{\diamond} \leq \|\mathcal{A}\|_{\diamond}\|\mathcal{B}\|_{\diamond}$, the fact that $\Theta_+^{(M)}\mathcal{E}\Theta_{in}$ is a channel and that $\Theta_+^{(k)}\mathcal{U}MeasPrep_{M,k}\Theta_+^{(M)} = \mathcal{U}MeasPrep_{M,k}$ and $\Theta_+^{(k)}\text{Tr}_{M-k}\Theta_+^{(M)} = \text{Tr}_{M-k}$. The two bounds above prove Eq. (19). Eq. (20) then follows immediately from the relation $\log(1+x) \leq x$. \square

Since the input quantum information has to be spread uniformly over a large number of receivers, a finite group of $k \ll M$ receivers can only access a vanishing amount of information. This fact holds even if the overall channel \mathcal{E} is unitary (for example, if \mathcal{E} is the identity channel from a super-user holding all input systems to M users, each of them receiving one output system).

3 Conclusions

In this paper we have seen that the standard finite quantum de Finetti theorems can be naturally rephrased as theorems about the diamond-norm distance between the optimal measure-and-prepare channel from M to k copies and the trace channel Tr_{M-k} . The working principle of the theorems appears to be the simple fact that estimation and re-preparation from M to k copies becomes almost ideal whenever M is large with respect to k . This idea suggests that similar approximation theorems could be obtained from other measure-and-prepare protocols based on estimation, where the input is given by M copies of some state $|\psi_x\rangle, x \in X$ and the goal is to produce k approximate copies. In this case, one can

expect to obtain approximation theorems for multipartite quantum states in the linear span of the projectors $|\psi_x\rangle\langle\psi_x|^{\otimes M}$. The exploration of such generalizations is an interesting direction of future research.

Acknowledgements. I would like to thank D. Gottesman, R. Spekkens, I. Marvian, and A. Harrow for stimulating questions that helped me to improve the presentation. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

Appendix

The Appendix is devoted to the derivation of Eq. (6). To this purpose we will use the fact that every operator $\rho \in \text{Lin}((\mathcal{H}^{\otimes M})_+)$ can be written as a linear combination of the rank-one projectors $|\psi\rangle\langle\psi|^{\otimes M}$. An easy proof of this fact is given as follows: Let us write $|\psi\rangle = \sum_{k=1}^d \psi_k |k\rangle$. Then, we have (cf. Eq. 2 of Ref. [2])

$$|\psi\rangle^{\otimes M} = \sum_{\mathbf{n} \in \mathcal{P}_{M,d}} \psi_1^{n_1} \dots \psi_d^{n_d} \sqrt{\frac{M!}{n_1! \dots n_d!}} |\mathbf{n}\rangle,$$

and also

$$\frac{1}{M!} \left(\prod_{k=1}^d \frac{1}{\sqrt{m_k!}} \frac{\partial^{m_k}}{\partial \psi_k^{m_k}} \right) \left(\prod_{l=1}^d \frac{1}{\sqrt{n_l!}} \frac{\partial^{n_l}}{\partial \psi_k^{*n_l}} \right) |\psi\rangle\langle\psi|^{\otimes M} \Big|_{\psi=0} = |\mathbf{m}\rangle\langle\mathbf{n}|,$$

where the coefficients $\{\psi_k\}_{k=1}^d$ and their complex conjugates $\{\psi_k^*\}_{k=1}^d$ are treated as independent variables. This means that the operators $|\mathbf{m}\rangle\langle\mathbf{n}|$ are in the linear span of the projectors $|\psi\rangle\langle\psi|^{\otimes M}$ (indeed, the derivatives are limits of linear combinations, and, since we are in finite dimensions, any linear span is a closed set, containing all its limit points). Since the operators $\{|\mathbf{m}\rangle\langle\mathbf{n}|\}_{\mathbf{m}, \mathbf{n} \in \mathcal{P}_{M,d}}$ span $\text{Lin}((\mathcal{H}^{\otimes M})_+)$, the projectors $|\psi\rangle\langle\psi|^{\otimes M}$ also do. Note that the same conclusion would be obtained, through a lengthier calculation, by taking all possible derivatives with respect to the real parts $\{\text{Re}(\psi_k)\}_{k=1}^d$ and the imaginary parts $\{\text{Im}(\psi_k)\}_{k=1}^d$, instead of the derivatives with respect to the coefficients $\{\psi_k\}_{k=1}^d$ and their complex conjugates $\{\psi_k^*\}_{k=1}^d$.

Due to the above discussion, to prove Eq. (6) it is enough to characterize the action of $\mathcal{UMeasPrep}_{M,k}$ on a generic projector $|\psi\rangle\langle\psi|^{\otimes M}$.

Moreover, since the choice of the basis $\{|1\rangle, |2\rangle, \dots, |d\rangle\}$ is arbitrary, for given $|\psi\rangle$ we can choose $|1\rangle = |\psi\rangle$. Then, Eq. (5) gives

$$\mathcal{U}MeasPrep_{M,k}(|1\rangle\langle 1|^{\otimes M}) = \frac{d_+^{(M)}}{d_+^{(M+k)}} \sum_{\mathbf{m}, \mathbf{n} \in \mathcal{P}_{k,d}} \alpha_{\mathbf{m}, \mathbf{n}} |\mathbf{m}\rangle \langle \mathbf{n}|$$

with $\alpha_{\mathbf{m}, \mathbf{n}} = \langle 1|^{\otimes M} \langle \mathbf{m} | P_+^{(M+k)} | 1 \rangle^{\otimes M} | \mathbf{n} \rangle$. Using the relation

$$P_+^{(M+k)} = \frac{1}{(M+k)!} \sum_{\pi \in S_{M+k}} U_\pi^{(M+k)}$$

and Eq. (2) with the substitution $M \rightarrow k$, we obtain $\alpha_{\mathbf{m}, \mathbf{n}} = \frac{k!(M+n_1)!}{(M+k)!n_1!} \delta_{\mathbf{m}, \mathbf{n}}$, and, therefore,

$$\mathcal{U}MeasPrep_{M,k}(|1\rangle\langle 1|^{\otimes M}) = \frac{d_+^{(M)}}{d_+^{(M+k)}} \binom{M+k}{k}^{-1} \sum_{\mathbf{n} \in \mathcal{P}_{k,d}} \binom{M+n_1}{M} |\mathbf{n}\rangle \langle \mathbf{n}|. \quad (21)$$

Using again Eq. (2) with the substitution $M \rightarrow k$ we get the chain of equalities

$$\begin{aligned} & \sum_{\mathbf{n} \in \mathcal{P}_{k,d}} \binom{M+n_1}{M} |\mathbf{n}\rangle \langle \mathbf{n}| = \\ & = \sum_{\mathbf{n} \in \mathcal{P}_{k,d}} \left(\frac{\binom{M+n_1}{M}}{k!n_1! \dots n_d!} \sum_{\pi, \sigma \in S_k} U_\pi^{(k)} (|1\rangle\langle 1|^{\otimes n_1} \otimes \dots \otimes |d\rangle\langle d|^{\otimes n_d}) U_\sigma^{(k)} \right) \\ & = \sum_{n_1=0}^k \frac{\binom{M+n_1}{M}}{k!n_1!(k-n_1)!} \sum_{\pi, \sigma \in S_k} U_\pi^{(k)} (|1\rangle\langle 1|^{\otimes n_1} \otimes (I - |1\rangle\langle 1|)^{\otimes (k-n_1)}) U_\sigma^{(k)} \\ & = \sum_{n_1=0}^k \sum_{j=0}^{k-n_1} \frac{(-1)^j \binom{M+n_1}{M} \binom{k-n_1}{j}}{k!n_1!(k-n_1)!} \sum_{\pi, \sigma \in S_k} U_\pi^{(k)} (|1\rangle\langle 1|^{\otimes (n_1+j)} \otimes I^{\otimes (k-n_1-j)}) U_\sigma^{(k)}. \end{aligned}$$

Defining $s = n_1 + j$, the chain can be continued as

$$\begin{aligned}
& \sum_{\mathbf{n} \in \mathcal{P}_{k,d}} \binom{M+n_1}{M} |\mathbf{n}\rangle \langle \mathbf{n}| = \\
&= \sum_{n_1=0}^k \sum_{s=n_1}^k \frac{(-1)^{s-n_1} \binom{M+n_1}{M} \binom{k-n_1}{s-n_1}}{k!n_1!(k-n_1)!} \sum_{\pi, \sigma \in S_k} U_\pi^{(k)} \left(|1\rangle \langle 1|^{\otimes s} \otimes I^{\otimes(k-s)} \right) U_\sigma^{(k)} \\
&= \sum_{s=0}^k \sum_{n_1=0}^s (-1)^{s-n_1} \binom{M+n_1}{M} \binom{k}{s} \binom{s}{n_1} P_+^{(k)} \left(|1\rangle \langle 1|^{\otimes s} \otimes I^{\otimes(k-s)} \right) P_+^{(k)}
\end{aligned}$$

Finally, we can use the combinatorial identity (see proof below)

$$\beta_s := \sum_{n=0}^s (-1)^{s-n} \binom{s}{n} \binom{M+n}{M} = \binom{M}{s} \quad (22)$$

to obtain

$$\sum_{\mathbf{n} \in \mathcal{P}_{k,d}} \binom{M+n_1}{M} |\mathbf{n}\rangle \langle \mathbf{n}| = \sum_{s=0}^k \binom{k}{s} \binom{M}{s} P_+^{(k)} \left(|1\rangle \langle 1|^{\otimes s} \otimes I^{\otimes(k-s)} \right) P_+^{(k)}. \quad (23)$$

Since $\binom{M}{s} = 0$ whenever $s > M$, the sum is in fact a sum from 0 to $\min\{M, k\}$. Combining Eqs. (21), (23), and (7) we obtain the expression

$$\begin{aligned}
\mathcal{U}MeasPrep_{M,k}(|1\rangle \langle 1|^{\otimes M}) &= \sum_{s=0}^{\min\{k,M\}} \frac{d_+^{(M)} \binom{k}{s} \binom{M}{s}}{d_+^{(M+k)} \binom{M+k}{k}} P_+^{(k)} \left(|1\rangle \langle 1|^{\otimes s} \otimes I^{\otimes(k-s)} \right) P_+^{(k)} \\
&= \sum_{s=0}^{\min\{k,M\}} \frac{\binom{M}{s} \binom{d+k-1}{k-s}}{\binom{d+M+k-1}{k}} \mathcal{U}Clon_{s,k} \left(|1\rangle \langle 1|^{\otimes s} \right),
\end{aligned}$$

which holds for arbitrary M and k , and for an arbitrary vector $|1\rangle = |\psi\rangle$. Hence, we have obtained Eq. (6).

Regarding the combinatorial identity of Eq. (22), it can be proved as follows: First, using Chu-Vandermonde formula (Eq. (8)) one obtains

$$\beta_s = \sum_{n=0}^s \sum_{l=0}^M (-1)^{s-n} \binom{s}{n} \binom{s+n}{l} \binom{M-s}{M-l}$$

Then, Klee's identity

(Proposition 1.1 of Ref. [24]) yields $\beta_s = \sum_{l=0}^M \binom{s}{l-s} \binom{M-s}{M-l} = \sum_{l'=0}^{M-s} \binom{s}{l'} \binom{M-s}{M-s-l'}$. Finally, the expression $\beta_s = \binom{M}{s}$ follows by applying Chu-Vandermonde formula again.

References

1. N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
2. R. F. Werner, Phys. Rev. A **58**, 1827 (1998).
3. D. Bruß, A. Ekert and C. Macchiavello, Phys. Rev. Lett. **81**, 2598 (1998).
4. M. Keyl and R. F. Werner, J. Math. Phys. **40**, 3283 (1999).
5. D. Bruß, M. Cinchetti, G. M. D'Ariano, and C. Macchiavello, Phys. Rev. A **62**, 12302 (2000).
6. G. M. D'Ariano and C. Macchiavello, Phys. Rev. A **67**, 042306 (2003).
7. J. Bae and A. Acín, Phys. Rev. Lett. **97**, 030402 (2006).
8. G. Chiribella and G. M. D'Ariano, Phys. Rev. Lett. **97**, 250503 (2006).
9. M. Keyl, Problem 22 of the list <http://www.imaph.tu-bs.de/qi/problems/>.
10. M. Christandl, R. Koenig, G. Mitchison, and R. Renner, Comm. Math. Phys. **273**, 473 (2007).
11. R. Renner, Nature Physics **3**, 645 (2007).
12. R. Koenig and G. Mitchison, J. Math. Phys. **50**, 012105 (2009).
13. C. M. Caves, C. A. Fuchs, R. Schack, J. Math. Phys. **43**, 4537 (2002).
14. B. de Finetti, Theory of Probability (Wiley, New York, 1990).
15. S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
16. V. Scarani, S. Iblisdir, N. Gisin, A. Acín, Rev. Mod. Phys. **77**, 1225 (2005).
17. D. Aharonov, A. Kitaev, and N. Nisan. Quantum Circuits with Mixed States. In Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC), ACM (1998).
18. V. I. Paulsen, *Completely bounded maps and dilations*, Longman Scientific and Technical (1986).
19. R. Askey, *Orthogonal polynomials and special functions*, CBMS-NSF Regional Conference Series in Applied Mathematics, **21**, Philadelphia, PA (1975).
20. J. Yard, P. Hayden and I. Devetak, arXiv:quant-ph/0603098v1.
21. G. Chiribella, G. M. D'Ariano, and P. Perinotti, J. Math Phys. **50**, 042101 (2009).
22. D. Leung and G. Smith, Comm. Math. Phys. **292**, 201 (2009).
23. A. S. Holevo and R. F. Werner, Phys. Rev. A **3**, 32312 (2001).
24. V. Klee, Canad. J. Math. **16**, 517 (1963).